Final Report of the

Joint Working Group of the

Social Sciences and Humanities Research Council

and

Statistics Canada

on the Advancement of Research using Social Statistics

December 1998

Executive Summary

Canada's economy and society are in a period of rapid and difficult change. Timely and objective analysis of economic and social conditions is required to understand this transformation, to provide a basis for broad and informed debate on public policy, and to establish a foundation for intelligent policy formation. The need is particularly acute because Canada's social policy has not kept pace with the dramatic changes in its economic policy over the past two decades. Governments at all levels have acknowledged the importance of redesigning social policy so that it meets the needs of all Canadians, and leads us towards more civil and economically sustainable communities.

In one respect, Canada is well-equipped to meet these needs. We now have a number of excellent and timely social surveys covering a variety of topics. Theoretical developments and advances in research design have led to the creation of longitudinal surveys which track individuals over extended periods of time. These new research tools provide information on the dynamics of poverty, the effectiveness of training programs, the consequences of job loss, the influence of childhood experiences, and several other topics pertinent to redesigning social policy. Together, these form the basis for establishing a well-integrated system of "social statistics", a term we use to encompass information describing a wide range of human activity and the social, economic, educational, and cultural features that affect our daily lives.

However, as a nation we have very little capacity to conduct social policy research, evaluate social programs, or monitor progress towards achieving social aims. The Federal Government recently acknowledged the need to strengthen its research capacity and established the Policy Research Initiative to recommend and oversee the implementation of an interdepartmental research agenda. Similarly, provincial governments, the private sector, and non-government organisations recognized this need and have attempted to revitalize the policy research community.

There are at least three significant barriers that need to be overcome if we are to develop our research capacity in social statistics. The first is the sheer lack of trained researchers. More than a decade of Federal and Provincial Government restructuring and downsizing has significantly reduced the number of researchers working within government institutions. This has occurred during a period of decline in the training of statistics and research methods in social science departments of our universities, with the exception of economics. The second significant barrier is a lack of access to data. Paradoxically, the detail provided by new data sets, which makes them so valuable, prevents them from being made public, because of their potential to identify specific individuals. The Statistics Act sets out strict criteria for maintaining confidentiality, and Statistics Canada necessarily has a very strong position on ensuring the law is followed. The third barrier is the very weak links between the work of social scientists and the potential users of the knowledge they generate. Even though there is a tremendous appetite for social statistics about education, employment, health, literacy, and other pertinent social issues, many of the important findings of social scientific research have not been adequately conveyed to the policy community, or to a wider public through the popular media.

The mandate of our working group, convened jointly by Dr. Ivan Fellegi, the Chief Statistician, and Dr. Marc Renaud, President of the Social Sciences and Humanities Research Council (SSHRC), is to make proposals to encourage quantitative research on major social and economic issues using large-scale data. This report recommends the funding of three components, each designed primarily to surmount the barriers described above. Together, they comprise a **Social Statistics Research System**, which would complement the integrated system of social statistical data for which Canada is already considered a world leader.

The first component is aimed at increasing the number of researchers engaged in quantitative research on social and economic issues. It includes three separate proposals: **Research and Training Groups** that would bring together researchers from different disciplines and institutions to conduct quantitative research and training in priority areas; a **Training Program**, including a summer school, that would provide specialized training in advanced statistical methods complementary to apprenticeships and graduate programs and support for data librarians to increase the use of social statistics in undergraduate training; and a **Fellowship Program**, including M.A., PhD and Postdoctoral Fellowships aimed at providing support for young researchers pursuing careers in social statistics, as well as Senior Fellowships that would enable some of Canada's leading social scientists to devote more of their time to research, and provide leadership in the training of the next generation of researchers.

The second component entails two proposals: **Research Data Centres** and remote access capabilities

that would provide access to detailed micro-data, while maintaining the strict rules for the preservation of confidentiality required of Statistics Canada under the Statistics Act; and support for the enhancement and expansion of the activities of the **Data Liberation Initiative**.

The third component is the development of a **Social Statistics Communication Program** that would implement a communications strategy to inform and build the public constituency for quantitative social science research. A key component of this program are **Research Forums** that would support research networks, provide an arena for the presentation of research findings, and enhance communication among researchers, the policy community, and the media.

The report also recommends that Statistics Canada and SSHRC negotiate a memorandum of understanding that defines the goals and organization of the proposed Social Statistics Research System. A co-ordination structure is also recommended, so as to take advantage of opportunities to strengthen networks and increase synergy among researchers.

Table of Contents

Page

I. Introduction 1

II. Building Research Capacity 3

A. Research and Training Groups 3
B. Improving Skills: Training Program and Summer School 5
C. Fellowship Programs 6
MA, PhD and Postdoctoral Fellowships 6
Senior Fellowships in Social Statistics 6

III. Access to Data 7

A. Research Data Centres 7
Remote Access 9
Seeking Approval from Respondents to Share the Data with Data Centre Users 9
B. Archiving and Documenting Data 10
C. The Data Liberation Initiative 10

IV. Communicating Research Findings 11

A. Research Forums 12
B. Social Statistics Communications Program 13

V. Co-ordination, Costs and Timetable 14

A. Co-ordinating the Canadian Social Statistics Research System 14
B. Costs and Timetable 15

Appendix A – Some Priority Areas for Research in Social Statistics 16

Appendix B – Curriculum for a Summer School Program 21

Appendix C – Research Data Centres 23

Appendix D – The Development of a Remote Access Capability at Statistics Canada 28

Appendix E – A Model for Research Forums in Social Statistics: the Canadian Employment Research Forum 29

Appendix F – Membership of the Joint Working Group 30

I. Introduction

Over the last decades, Canada has experienced - and continues to experience - dramatic social, economic, and technological changes. Our economic policy has responded to and adjusted to these changes for a number of reasons. In particular, however, the adjustment has been in response both to pressures stemming from global international markets, and to the need to abate inflation and reduce the national debt. These changes in economic policy have been accompanied by major changes in the labour market and in family structure. There is a general sense among many Canadians that the major problems we face are not economic, but social. Governments at all levels have acknowledged the need to redesign our social policy so that it fits better with our current economic policy.

To meet this need we require research on a wide range of social, economic, educational, and cultural issues. The renovation of social policy must build on a basic understanding of the life course, and the complexity of social relationships at different levels, such as families, neighbourhoods, and communities. Conducting such research requires a well-integrated system of social statistical surveys, together with the capacity to analyse the data.

Beginning with data from the census and surveys on trade, finances, prices, the labour force and other areas, Statistics Canada has developed a national statistical system. Historically its aim has been to provide aggregate indicators of current economic and social conditions, including descriptive statistics on economic growth, health, education, justice, productivity, the labour force, and so on. During the 1970s, two major changes shaped its evolution. First, research in the social sciences demonstrated that an understanding of many social phenomena - such as the nature of criminal activity and victimization, or the effects of poverty - required separate, focused surveys. Second, the policy community recognized the importance of research that could help us understand how particular life events, combined with peoples' habits and lifestyles affected their long-term social outcomes. Research began to stress the importance of *social context* – the families, neighbourhoods, schools, and organizations in which people live and work – in shaping and constraining individuals' actions. But the descriptive data available from cross-sectional surveys were inadequate for monitoring changes in social outcomes, or understanding the causal mechanisms that led to desirable outcomes. This required longitudinal surveys, in which data were collected from the same sample of respondents on at least two, and preferably more occasions.

Statistics Canada responded to this challenge by instituting a new generation of social surveys during the 1980s and early 1990s. They include, among others, the General Social Survey (GSS), the Graduate Follow-up Surveys, the Survey of Labour and Income Dynamics (SLID), the National Longitudinal Survey of Children and Youth (NLSCY), the National Population Health Survey (NPHS), and the Displaced Worker Survey (DWS). Over the next few years, additional major new surveys are planned, including the longitudinal Workplace and Employee Survey (WES), the longitudinal Youth in Transition Survey (YITS) and the Survey of Financial Security (SFS). Meanwhile, Statistics Canada also strengthened its efforts in collecting administrative data and taxation data, and its capabilities to link data from various sources. Canadian researchers also played a major role in the development and administration of international surveys, such as the International Adult Literacy Study (IALS) and the Third International Study of Mathematics and to Science (TISMS). Several other surveys are currently being developed, and efforts are being made to integrate the survey data so that they can better characterize Canadian life. Together, these surveys have taken the national statistical system in a new direction, and have helped distinguish Canada as a world leader in the collection of social statistical data.

Despite this great strength, Canada has relatively little capacity to analyse the data from these surveys. The most difficult barrier to surmount is that there are simply too few researchers engaged in quantitative research. The problem is especially acute in areas requiring advanced statistical methods. The capacity to conduct research within government institutions has declined over the past two decades, mainly due to restructuring and downsizing. Although these institutions appreciate the need to bring analyses of social statistics to bear on public policy debates, they are not strongly enough connected to the universities, where the majority of social scientists conduct their research, and where the training of new researchers occurs. Within the universities, training in statistics and research methods has also declined severely during this period. There are now very few faculty who teach courses in advanced statistical methods, and in most social science fields, the course requirements for M.A. or PhD degrees no longer include formal training in statistics or quantitative research methods. The notable exception is economics, but in this field research

capacity has been and continues to be eroded by the migration of qualified researchers to the United States.

This problem is exacerbated by the fact that the data generated by the new surveys are relatively complex: the surveys are usually longitudinal, and have a multilevel structure, such as students nested within schools, or workers within firms. Thus their analysis requires powerful computing and statistical techniques. There have been tremendous theoretical developments in this area, but there are very few social scientists in Canada who have training in their application.

Another significant barrier is access to data. In the past, much useful research was conducted on public use data files made available to university and other researchers by Statistics Canada, especially after the launching of the Data Liberation Initiative (DLI) by the academic community, with support from Statistics Canada, the Humanities and Social Sciences Federation of Canada, and SSHRC. A difficult problem is posed by the need to make detailed micro-data – the exact responses to questions in surveys – available to researchers, without compromising the confidentiality promised to survey respondents. The powerful statistical techniques which are appropriate for the analysis of multilevel, longitudinal data cannot be performed with aggregate data; access to micro-data is essential. Moreover, even simple descriptive problems often require access to micro-data. Part of the problem, ironically, is that the detail on individuals provided by longitudinal micro-data could make it possible for a researcher to identify specific individuals. This would violate requirements to maintain the confidentiality of individuals' responses, as set out in the Statistics Act.

The third barrier to the development of our research capacity in the social sciences concerns communication. Many social scientists necessarily devote their energy to addressing narrowly defined research questions that will advance their academic discipline. Success in academic careers depends to a large extent on publication in scholarly journals, and usually there are relatively few incentives or resources available for preparing more popular pieces that would convey research findings to a broad audience. Moreover, the demands of an academic career to excel in research and teaching make it difficult for researchers to forge strong links with the policy community, and often the time span of academic research is too long to meet the immediate needs of policy-makers. Consequently, the transfer of knowledge from research to policy and practice is not as rapid or as strong as it could be. Perhaps for the same reasons, social scientists have not been particularly successful in popularizing their findings through the public media.

However, in the longer term, the strengthening of social science research requires widespread public recognition of the benefits of social policy research. We need to demonstrate that empirical research provides a basis for informing the development of public policy, evaluating social programs, and monitoring our progress towards achieving social aims. There is no substitute for a specialized dialogue among researchers and policy-makers. As this will not occur spontaneously, an organized effort is required to increase public awareness of research findings.

A Joint Working Group of the Social Sciences and Humanities Research Council (SSHRC) and Statistics Canada was convened by Dr. Ivan Fellegi, the Chief Statistician, and Dr. Marc Renaud, President of SSHRC. Its mandate is to present a set of proposals to encourage quantitative research on major social and economic issues using large-scale data. The effort of the Group takes place in the context of other initiatives aimed at improving research using social statistics. The Federal Government recently established the Policy Research Initiative (PRI), with a mandate to recommend and oversee the implementation of an interdepartmental research agenda designed to remedy gaps in the knowledge required for policy development in the Federal Government. The PRI has partnered with SSHRC to undertake a project on societal trends, with the objective of gaining insight into major social changes, to understand their implications for policy research, and to identify key areas for future research. In addition, SSHRC is developing programs of strategic research in the areas of social cohesion, the knowledge-based economy and society, and population health. These are also priority research areas for the PRI.

This report recommends the funding of three components, designed primarily to increase the number of researchers engaged in quantitative social science research, improve access to micro-data without compromising confidentiality, and inform and build the public constituency for social statistics. Together, these components comprise a **Social Statistics Research System**, which would complement the national statistical system which is already well developed. These components are described separately in the next three sections of this report. The final section sets out a time-line and estimated budget for implementing the proposals, and recommends a process for co-ordinating the inter-related activities that would follow.

II. Building Research Capacity

The first component of our proposed Social Statistics Research System aims to increase the number of researchers engaged in quantitative social science research, and to improve the skills of those already working in this field. It contains three separate proposals: **Research and Training Groups**, a **Training Program**, including a summer school, and a **Fellowships Program** that includes M.A., PhD, Postdoctoral, and Senior Research Fellowships. These are described below.

A. Research and Training Groups

> We recommend establishing a program of Research and Training Groups with four interrelated functions: the conduct of quantitative social research; the training of researchers; the broad dissemination of research findings; and the provision of feedback about the usability of data sources. We envision medium-size research groups, preferably cross-university and cross-disciplinary, working on major research problems, focused on data from a number of different sources. The program for Research and Training Groups would have an adjudicating committee, including researchers from abroad, that would implement a strategic, multi-stage, and pro-active adjudication process.

Central to the development of a strong and diversified body of research is a community of scholars who work together and encourage and criticize each others' work. Research communities that focus on particular research issues are especially critical for graduate students and researchers beginning their careers, as they enable them to accumulate experience and judgement which is not easily learned from books or through university courses. SSHRC has experience with a variety of different kinds of organizational forms, and it has found that the most productive environments are those where researchers meet face-to-face and work on a common intellectual theme.

The centrepiece of our initiative, therefore, is a program to support research and training groups. We envisage middle-sized groups, typically with 8 to 15 researchers. This number provides a critical mass, while avoiding the coordination problems of larger groups. Each Research and Training Group would focus on a topic that is sufficiently broad to merit a co-ordinated program of research, training, and dissemination activities; and sufficiently important to attract a significant number of researchers. Ideally the Groups would be multi-disciplinary, and entail partnerships among disciplines within a university, and between two or more universities. These Groups would focus their efforts on research pertaining to a particular theme. **Appendix A** (see p. 16) of this report describes some promising themes where there is already an abundance of available data.

There are already social research groups in Canada, in areas such as aging, education, immigration, and child development, with funding through SSHRC's Major Collaborative Research Initiatives (MCRI) and National Centres of Excellence. The proposed Research and Training Groups, however, would focus on quantitative social science research, and place much more emphasis on developing the skills of new researchers and enhancing the skills of an entire research team. Good research teams provide valuable informal learning as a by-product of conversations, meetings, and formal seminars, but the Research and Training Groups would be required to address the need for training more explicitly. They would identify

training needs and describe concrete steps to address them. They might create and improve formal courses, or prepare background research papers on research methods or aspects of survey design. Efforts to establish apprenticeships that would bridge the gap between formal graduate courses and applied research would be of primary importance.

The Groups would also be expected to develop and carry out plans for the dissemination of their research results, not only to specialized audiences of academics, but to policy makers and the general public. This would involve more than an episodic flurry of communications activity; it would entail sustained efforts at reaching out, culminating, in several cases, in the formation of partnerships with groups outside academia. Finally, these Research and Training Groups would provide feedback to those collecting data, in particular to Statistics Canada, about the quality of data being collected, its relevance to central research problems, and its limitations. We expect that this would improve future data collection efforts, and sustain researchers' commitments to this activity.

Research and Training Groups that work in similar or related areas would, over time, become involved in virtual networks – regionally, nationally, and internationally. These networks would allow researchers who were not geographically close to interact with the Groups, or even to become members. Research and Training Groups would be encouraged to pursue their training, dissemination, and feedback objectives by putting to use other components of the Canadian Social Statistics Research System that we propose: the Summer School and the specialized workshops to be offered by the Training Centre, the Research Forums, which would offer the opportunity of meeting with a variety of research users, and the Communication activities to be developed at SSHRC.

We expect that most of the Research and Training Groups would focus on the analysis of large-scale Statistics Canada surveys. SSHRC and Statistics Canada would encourage research that uses complex longitudinal and multi-level data, and comparative research, especially using multi-country databases. Research and Training Groups might also plan and conduct their own surveys, either free-standing, such as studies of Canadian elections, or surveys that involved Canadian participation in an international social survey. Conducting such surveys outside of Statistics Canada, or in partnership with it, would increase opportunities for training students how to plan, design, and carry out a social survey. This would benefit Statistics Canada and the survey research industry.

Applications for Research and Training Groups would be peer reviewed, with successful applications funded for an initial period of 5 years, and renewable. Typically, the project budgets would support student research assistants and postdoctoral fellows, computers and software, travel, and in some cases, data collection and archiving. Funding could also be granted to support non-student research professionals; in particular these Groups would likely require professionals capable of preparing and documenting the complex datafiles required to conduct sophisticated quantitative analysis, of providing ongoing support to graduate students who undertake data analysis, and of solving problems with computers and software.

We recommend that the Research and Training Groups program have a single, separate adjudicating committee, with international members who have expert knowledge of quantitative social science research, including policy-related research. The committee would be charged with developing an adjudication process that was comprehensive and formative; that is, one that encouraged the development of new proposals and links with other groups and the policy community. We feel a separate adjudication committee is necessary as it would be able to understand the special requirements of quantitative social sciences research, and balance the criteria, preferences and objectives to be served by the Research and Training Groups. It would also serve to situate the activity of the Groups within the larger context of the Canadian Social Statistics Research System, and attend to the on-going development of Groups, especially when they are in their early stages of their development.

B. Improving Skills: Training Program and Summer School

> We recommend the establishment of a Training Program, including a national Summer School, covering topics in quantitative methods and the use of Canadian micro-data. The program would serve policy analysts, university researchers, and graduate and undergraduate students. It would also provide support to data librarians.

To complement the Research and Training Groups, there is a need for a program of short courses in social data analysis, similar to the American and British efforts at Ann Arbor and Essex, but more decentralized and adapted to Canadian data and research concerns. The goal is not to offer a substitute for university courses at the graduate and undergraduate levels, but to provide a series of intermediate-level courses in data analysis, and more advanced courses pertaining to the analysis of longitudinal and multilevel data. These courses would stress the application of statistical techniques, and would use Canadian micro-data.

A Training Program of this nature is required for several reasons. One is that students at the senior undergraduate and graduate levels who take training in statistics are often unable to make the leap from the theoretical training offered in university courses to applying the techniques to real data. In many cases, the examples offered in textbooks pertain to social issues in the U.S., and employ small, contrived data sets. Many students need some basic tutoring on how to manage data, such as how to retrieve data from a CD-ROM, merge files, select cases, and sort data. They also tend to need instruction on rather simple topics which are seldom covered in their training, such as scaling variables, creating composite variables, handling missing data, and using design weights. Some basic courses that used one or two Canadian data sets could generate considerable enthusiasm for pursuing a career in social statistics. Also, when students take such courses alongside other university researchers and policy analysts, they learn first-hand about the kinds of problems that the analysis of micro-data can address. A considerable amount of incidental learning occurs, and important working relationships are formed, many of them inter-disciplinary.

Researchers who are at an advanced level also require training in the more complex statistical methods appropriate for analysing longitudinal and multilevel data sets. This is essential if we are to exploit the richness of the recently developed surveys. The audience for these courses would be mainly university researchers, and policy analysts working in government agencies and other organizations. These courses also need to be provided centrally, because there is not usually a critical mass of researchers wishing to pursue such advanced training at any one university.

Finally, because data librarians play an essential role in facilitating access to and use of the existing databases, efforts should be pursued to increase their training. This would contribute to undergraduate education as well, because data librarians can help professors use Canadian data in their courses, and assist students in getting started on projects that require some data analysis.

The logical centrepiece of an effort to improve training in quantitative data analysis in Canada is a summer program, modelled somewhat after the ongoing and successful summer programs held at Michigan and Essex. The idea is to offer a series of courses in one or more locations, normally over a three- to four-week period. Ideas for a curriculum are provided in **Appendix B** (see p. 21). However, this program would differ from the UK and U.S. programs in three important respects: first, it would focus on Canadian social issues and use Canadian data; second, it would provide a social and intellectual environment for graduate students, faculty members and public sector researchers with an interest in quantitative research on Canadian issues; third, it would provide a means for follow-up activities, such as establishing an electronic network and the exchange of research papers. It would also give researchers an opportunity to establish contacts with personnel at Statistics Canada who are directly involved in managing particular data bases.

There is no reason to restrict these activities to the Summer. The Training Centre should provide a platform for related educational and consultative activities throughout the year. These could include workshops to develop plans for new surveys, seminars to support research on a given topic or with a given data-set, and workshops on statistical techniques and survey methods. Departments, foundations, businesses, unions, and NGOs could use the facility for consultations and seminars on pressing research questions. Activities of this kind already take place on an *ad hoc* basis, but a Centre with a mandate to

facilitate such training would dramatically reduce the effort presently required for researchers to conduct such training programs.

We propose to begin with a program organized by Statistics Canada and using their teaching facilities in Ottawa. Aside from providing a good, ready-made teaching environment, Statistics Canada's ability to operate both in French and in English is extremely attractive. From the beginning, Statistics Canada would work closely with academics through an advisory committee. As the program becomes established, some of the courses could be taught at various universities. In the longer term, the program itself could be moved to a university base in one or more universities, while retaining a partnership with Statistics Canada. Also, there should be an effort to reach an accreditation agreement with some universities, as this would likely be helpful for many students.

We recommend that funding be provided to initiate the program. Funding is required for the salary of a director, office expenses, travel, instructors, and travel and honoraria of the advisory committee. The main source of funding for the Centre would be tuition fees and sponsorships for specific activities, such as workshops on a particular data-set. In the longer term, this program should strive towards becoming self-financing. Attracting graduate students to a summer program will require SSHRC support for travel, accommodation and tuition. Support of this kind might also be made available to junior faculty members. It would be appropriate to approach MRC and NSERC, who could support graduate students in areas that would benefit from this program.

C. Fellowship Programs

> We recommend that SSHRC establish a targeted program to support students and young researchers at the M.A., PhD and Postdoctoral levels in social statistics, and a program of Senior Research Fellowships in Social Statistics.

To develop a stronger research community in the longer term, we must ensure that young scholars are attracted to pursue a career in social statistics. The field needs a distinct and higher profile, visible to graduate students and post-docs as they decide on the orientation of their research. For this purpose, they should have access to specifically targeted financial support.

M.A., PhD and Postdoctoral Fellowships. The fellowships program would be distinct from the regular SSHRC fellowships programs, but its basic eligibility conditions, regulations and the size of awards would be the same. Applications would be restricted to quantitative social research, though the requirements for demonstrating this will vary according to the level of the award. M.A. candidates would be expected to take graduate courses that provide them with theoretical and analytic skills, and to write a review paper or thesis on a quantitative topic in social statistics. The application for PhD and Postdoctoral Fellowships would include a proposal to conduct a quantitative research project.

Given the targeted nature of this program, evaluation of these applications requires the establishment of a separate expert adjudication committee. Also, there should be an explicit commitment to achieving a gender balance. We therefore recommend that all applications for these fellowships be adjudicated directly at SSHRC.

Senior Fellowships in Social Statistics. The committee recommends that *two* Senior Fellowships be granted every other year to university scholars who are making a significant contribution to research and training in the area of Canadian social statistics. The Fellowships would be for $100,000 per year for a period of five years, and would require that recipients devote at least 80% of their time to research and training. Renewal would depend on the recipient maintaining a high level of scholarly research. The Fellowships would be prestigious, and would likely increase the momentum of the scholars' academic careers. These fellowships are modelled after those awarded by the Canadian Institute for Advanced Research (CIAR). The CIAR fellowships have been highly successful in keeping some of our best scholars in Canada, and enabling them to develop significant research programs.

We recommend the adjudication committee be the same as for the other fellowships proposed above.

Scholars could be nominated for a Senior Fellowship by their university, or they could apply directly to SSHRC. The application would entail a brief description of how their proposed research program would contribute to social statistics and to the training of young scholars. Adjudication would be based on the applicants' records of scholarly activity, and their proposed research and training programs.

Internships. We have also identified a need for more young researchers to obtain experience in policy research, in view, among other things, of eventually finding employment for their newly developed skills in quantitative social sciences. Although we see the immense potential benefit of a program of internships, we could not develop precise ideas in this respect in the absence of a more general program of internships at SSHRC. We can therefore only recommend that the new internship program proposed in the Council's Innovation Scenario take into account the special need for social statistics expertise identified in this report, and that appropriate resources be provided for it. Moreover, we think that in view of the development of the Social Statistics Research System being proposed here, Statistics Canada should more fully promote its own internship programs, and encourage other government departments to do the same.

III. Access to Data

If Canadians are to fully benefit from the substantial data resources that exist, researchers and analysts must have adequate access to these data. Research is required to convert data into usable information, and adequate access is necessary to facilitate and promote research. Data access has many dimensions. It can mean ensuring that users have adequate knowledge of and access to basic statistics from the data sources. It can also mean access to public use micro-data files required for more complex analysis, and at yet another level, access to the detailed unscreened (confidential) micro-data. Archiving is also an access issue, as data sources may not be accessible in the future if care is not taken to properly archive them when they are created.

The Data Liberation Initiative (DLI) significantly improved the access to public-use household survey data sources. The Joint Working Group viewed this initiative as extremely positive, and recommendations are made in this section regarding ways in which the initiative might be extended. Recommendations are also made regarding support for the archiving of data created by research groups and initiatives outside of Statistics Canada.

The major issue dealt with in this section, however, relates to the access by researchers to data which are confidential under the Statistics Act. The principal issue is how to create such access while ensuring that the respondents' confidentiality is protected as required under the Statistics Act. Confidentiality is the cornerstone of the statistical system. It is also true that in order to fully exploit the new data sources being created, researchers need to have some form of access to the detailed data. A number of alternative means of achieving these goals are discussed here, but the main proposal relates to the creation of Research Data Centres

A. Research Data Centres

> The committee recommends that SSHRC and Statistics Canada jointly create a national system of Research Data Centres, where researchers can access detailed micro-data for research purposes, while maintaining the confidentiality provisions of the Statistics Act.

Increasingly researchers require access to detailed micro-data to conduct research in many areas, including health, child development, income and economic security, labour adjustment and workplace change. Paradoxically, as the quality and complexity of the data have increased, access to these micro-data may decline, as a result of confidentiality requirements. Confidentiality protection is crucial. Not only is there a commitment of confidentiality of responses made to citizens when they provide information, but the security of data is vital to maintaining the public and political confidence required to carry on the broad range of data collection activities effectively.

Since the early 1970s, Statistics Canada has made micro-data (individual records) available while protecting confidentiality through the production of "public use" micro-data files for household surveys. To prevent the release of records in which the information from individual respondents can be identified, some data must be suppressed: this may involve the omission of some elements of the individuals' records or "collapsing" response categories, so it is not possible to identify respondents with unusual characteristics. But this solution is no longer adequate for two reasons. First, many of the new longitudinal surveys have a data structure that prevents public-use micro-data files from being created at all, for it becomes increasingly difficult to develop a micro-data file that maintains confidentiality when longitudinal data are available for a number of years. Second, researchers with more complex theories and more powerful statistical tools increasingly find that the suppression of information in traditional public use files significantly limits their analysis. If the data access issue is not addressed, not only will a number of new and innovative data sources be underutilized and valuable research lost, but researchers may increasingly turn to data from other countries, particularly from the U.S., where the access problem has been more adequately addressed. Findings from American research data are often not applicable to the Canadian situation.

One means of addressing this issue is to have researchers sworn in under the Statistics Act as "deemed employees" of Statistics Canada, allowing them to conduct research using the detailed micro-data files. The Statistics Act provides for such deemed employee status under conditions that are discussed later. Even after being sworn in under the Act, however, there is still the issue of physical access to the data. It is proposed that a number of locations be created across Canada where such access would be possible. These offices would be called **Research Data Centres**, and would be modelled on a similar program established by the U.S. Census Bureau. Following is an overview of the Research Data Centre proposal. The details of the proposal can be found in **Appendix C.** (see p. 23)

Ultimately, we envision several research data centres distributed across major urban centres and university campuses throughout the country. A centre would provide a secure physical location for the confidential data. Legally it would be an office of Statistics Canada, and a Statistics Canada employee would be on site. The centres would operate a research program administered by SSHRC and Statistics Canada. Researchers who had a proposal accepted under the program, and were sworn in under the Statistics Act, would have access to confidential micro-data files maintained at the Centres. A committee that operated under the auspices of SSHRC and Statistics Canada would carry out the adjudication of the program submissions, which would be on the basis of merit. The Selection and Review Committee(s) would consist largely of academic researchers, with some representation from other research communities and Statistics Canada. The research merits of the proposal would be paramount in the selection process (combined with resource issues). Researchers at universities, research institutes, government agencies, and other research organizations could submit proposals.

The Statistics Act governs access to confidential data, and its conditions must be incorporated into any program of Research Data Centres. Researchers may become deemed employees of Statistics Canada and access confidential data providing that they are "perform(ing) special services for the Minister (i.e. Statistics Canada)." Since Statistics Canada clearly does not have the human resources necessary to conduct all the research necessary to exploit the data sources, arranging for other researchers to provide such service is one means of achieving this goal. In this environment the research done would be similar to that which Statistics Canada itself would normally conduct. While this may sound restrictive, in fact Statistics Canada carries out a wide range of quantitative research; therefore, this is not likely to be an issue.

The output of the research program associated with the data centres would consist, in the first instance, of research papers produced by the researchers. The papers would undergo the normal peer review process, which would be managed by the Selection and Review Committee. It is anticipated that the vast majority would then be included in the Research Paper Series associated with the research program. Since the research program would be run jointly by SSHRC and Statistics Canada, these papers could not comment directly on policies or programs, as that is outside the mandate of Statistics Canada (Appendix C on p. 23 provides more detail.) Beyond this first stage, the researchers would be free to publish a revised paper (or the original research paper) in an academic journal, or any other publication. There would be no concern in this case as to whether the product contained policy comment; the authors would be free to include any material they deemed necessary. It is proposed that the copyright for the original research paper would be vested with the researchers. Statistics Canada would retain the right to publish the original research paper, and to vet all publications for confidentiality.

As long as the site had a level of security comparable to that maintained by Statistics Canada, a Research Data Centre could be located at a university or a non-governmental research institution. For convenience, however, some Centres could be attached to existing Statistics Canada regional offices. The advantage of non-Statistics Canada locations is that they would allow the institution serving as a home base to develop a very strong empirical research capacity. Empirically-oriented researchers would be attracted to an institution with a Data Centre, allowing the university or institute to develop a strong program. However, it may be appropriate to begin this program with a pilot project at a Statistics Canada regional office. It is proposed that Research Data Centres initially include surveys and administrative datafiles with the exception of the ones concerning businesses; issues arise in the access to the latter that would need to be addressed before they were included in the program.

There are at least two alternative funding approaches for research data centres. The first, modelled on the Data Liberation Initiative (DLI), is for the Centres to be entirely "block funded", with the costs shared by some combination of SSHRC, Statistics Canada, the university or institute at which the centre is located, and other organizations, such as government departments and research institutes. If, as we anticipate, primarily academic researchers would use the centres, the majority of the block funding would be SSHRC-based. In this model, there would be no direct cost to the researchers using the centres. The selection process, conducted by the selection and review committee, rather than applications for funding, would regulate access to the centres. A variant of this approach would be to have part of the annual cost of the Centres covered by block funding, and the remainder covered by fees paid by researchers using the centre. University-based researchers whose projects were approved by the committee would receive a SSHRC grant, distributed through the review committee to cover this cost; researchers from other organizations would pay similar fees. In any case, care should be taken to facilitate access to data for graduate students and post-docs.

Remote Access. In addition to deemed employees accessing confidential data at Research Data Centres, there are at least two complementary potential solutions to the data access problem. One is to develop a remote access capability in Statistics Canada. Researchers would submit their computer jobs to the agency, and they would be run on confidential micro-data files by Statistics Canada staff. The output would be vetted by Statistics Canada staff to ensure that confidential data were not released. This approach is complementary to, and not a substitute for, the Research Data centres. It would be useful for researchers with smaller projects, or those who were not physically close to the Research Data Centres. The remote access would not, of course, allow the flexibility and quick interaction required by many research projects that the Research Data Centres would provide. A Statistics Canada committee has been struck to develop a proposal in this area, and an outline of their proposal is in **Appendix D**. (see p. 28)

Seeking Approval from Respondents to Share the Data with Data Centre Users. The Statistics Act includes a provision for the sharing of data with users. If the respondent's permission is sought at the time of data collection, confidential micro-data may be shared with selected users. Data sharing would be possible with an Institute created by SSHRC, or some other body. Statistics Canada would then provide the raw micro-data (for respondents who agreed to share, most of whom do) to the Institute, which would of course agree under a contract arrangement to maintain confidentiality.

This approach is quite straightforward. The selection and vetting process would then be entirely in the hands of the Institute; Statistics Canada would play little role other than setting up the conditions to ensure data confidentiality. The shortcoming of this approach is that it can only be applied to data collected in the future, and to some data sets. The data sets that have been created to date could not fall under such an arrangement. For that reason, this Joint Working Group believes that the approach should be seriously considered for the future, but will not solve the current issues. In the longer run this approach is a potentially excellent solution to sharing selective data sets, and should be seriously considered once the initiatives in this report have reached some level of maturity.

B. Archiving and Documenting Data

> We recommend that SSHRC allow researchers who created data files before documentation and archiving was an allowable expense in SSHRC grants to apply for funding to conduct such documentation and archiving on the more important older data files. We also recommend that SSHRC and/or DLI consider the creation of a national archiving system for quantitative social science data sets created outside of Statistics Canada.

Organizations other than Statistics Canada often produce large quantitative research databases. SSHRC-funded initiatives can generate such files. The more important of these data sources should be archived for future use, and the data bases should be created and maintained so that they are readily accessible and transportable. If this is not done, researchers other than those initially involved with the data have no access to it, particularly some time after the file was created. There is currently a lack of support for archiving important research databases created some time ago. In SSHRC-funded projects involving the collection of data, the cost of providing documentation and creating files for dissemination purposes has become an allowable expense. The difficulty is with databases assembled without SSHRC support or with only partial support over a period of time and those that pre-date the funding allowances for archiving. Examples of such data files include The Aging in Manitoba data set, The Canadian Study of Health and Aging, and The Canadian Fertility Studies. Our understanding is that at this time, it is not possible to apply for funding to place the data in a form that can be disseminated. We therefore recommend that SSHRC expand one of its existing programs or develop a new program to provide for such funding. Applicants would have to make the case that the data are not presently in a form that allows them to be easily distributed or archived, provide plans for preparation of the data, a description of the final products and demonstrate that the data are of sufficient importance and interest that they will actually be used if archived and disseminated.

C. The Data Liberation Initiative

> We recommend that the Data Liberation Initiative be invited to submit proposals with a view to pursuing and extending its mission of making data and research information easily available for academic research and for undergraduate training; this could include the training of data librarians, and the extension of networks of co-operation among them and with undergraduate instructors, as well as the preparation, dissemination and exchange of teaching materials.

The Data Liberation Initiative (DLI) was launched a few years ago by the Humanities and Social Sciences Federation and Statistics Canada, together with a consortium of post-secondary institutions and federal departments. It has been an immense success, allowing researchers in universities and colleges much easier access to public use data files. Many more researchers are developing an interest in quantitative data. A new generation of data librarians have been appointed and trained in most institutions, and they have learned to work in close co-operation with one another to increase the availability of both data and the information required to analyze them. Prior to the DLI, data centres existed in very few institutions, and now they are being created in many more. Software has been developed to provide meta-data on available data sources, and to help with their extraction: students and relatively inexperienced researchers can now easily and rapidly bring quantitative data to bear on their analytical ideas; this nurtures their initial interest, and paves the way for further involvement with social statistics.

Throughout this period, many data librarians have also engaged in collaborations with social science instructors who teach undergraduate students the logic, methods, and practice of data analysis. Many of these instructors have incorporated in their teaching much more hands-on experience with data by students. Moreover, they can send their students, graduate as well as undergraduate, to the data librarians to get help in accessing the data and the accompanying documentation.

The Data Liberation Initiative now is under review as it approaches the end of its initial five-year phase. It is a free-standing organization, with its own Board and administration, and it is not the role of our Joint Working Group to make their case in their stead. This being said, we need to say how important it is that they be supported by Statistics Canada and SSHRC, morally as well as financially. The original mission of disseminating public use data to academic researchers should of course be preserved; but other roles that have been assumed by the DLI should be recognized, pursued and augmented. First, the training of data librarians, who have become indispensable intermediaries between data sources and users, should be expanded, and the standards of the profession should be sanctioned by employers and employees alike. Also, the networks of co-operation among them should be further extended and strengthened.

Second, the DLI is ideally placed to assume a larger and more systematic role in improving undergraduate education in methods and social statistics, in co-operation with post-secondary institutions, as well as with the Humanities and Social Sciences Federation, Statistics Canada, and SSHRC. It can promote exchanges among undergraduate instructors, and with data librarians, about improved ways of teaching, through written and electronic exchanges as well as in workshops, for instance at the annual Congress of the Social Sciences. It can also improve the availability of teaching materials for these instructors, in the form of raw or semi-processed data, documentation about data and concepts, articles and research reports, and suggestions for course assignments.

Third, the work of the DLI, centrally as well as in the local data centres and in networks, may be seen as the precursor for an eventual Canadian data archive. Such archives already exist in most advanced countries, and they have proved invaluable research instruments. While the creation of such an archive at the national level is not presently planned, that may change, and the work of the DLI in producing better meta-data and documentation is most likely to move us in that direction. Therefore, it should be strongly supported. Indeed, the proximity of data librarians to the users will help to keep the material readily usable by the latter. We can anticipate that university-based data centres will even succeed in preserving and disseminating the precious procedural information developed in the course of using quantitative data in research. In other words, a data archive might grow from the bottom up over the first few years of maturation of the Canadian Social Statistics Research System; SSHRC and Statistics Canada should pay attention to potential developments, and be prepared to lend them support.

IV. Communicating Research Findings

Ultimately, support for an active program of quantitative social analysis depends on public support and interest. Indeed, supplying relevant research and bringing it to the attention of people who could use it in debate and decision-making is key to increasing demand for such research. But at best, the process of recruitment and incentives around academic research does little to reward the abilities and temperament required to arouse public interest. With notable exceptions, policy-makers find contemporary social science either not directly relevant to their problems, or too narrowly focused. And researchers, in turn, are often reluctant to engage in work that appears unduly oriented towards short-term practicalities. The links between social science research and the media are also relatively weak, in spite of the tremendous appetite of the public for social statistics about education, employment, health, social assistance, aging and other social issues. In both cases, obstacles crystallise, in particular, around the dispersion of the information, the inaccessibility of the language, and differences in operational timetables.

While some progress has been achieved in certain research areas, we need to be more proactive in the promotion of dialogue between social scientists and the potential users of their findings. What is needed is a change in culture, which can only come about after a rather extended period of interaction; and such interaction will in turn require that opportunities be systematically created and exploited for this purpose. This is why we propose the creation of Research Forums, as well as the organization of a Social Statistics Communications Program at SSHRC.

A. Research Forums

We propose the establishment of a program for the support of research forums in order to bring together social researchers based in universities, governments and non-governmental research organizations with a broad range of policy analysts and decision-makers. The forums would sponsor conferences and related activities at which academic and government researchers would debate, and would benefit from each other's research, data, and experience.

Too often today, academic research on social and economic questions of interest to policymakers in federal, provincial and municipal governments is never seen by those potential audiences. At the same time, academic researchers remain unaware of emerging policy debates to which their expertise can usefully contribute. As a result, a huge pool of expertise for the greater understanding of Canadian social trends and policy issues remains untapped. To bridge these gaps, we propose a program of Research Forums.

Our model for Research Forums is the Canadian Employment Research Forum (CERF), which was formed in 1991 to foster Canadian research and bring together policy researchers in Ottawa and academic labour economists. As recently as ten years ago, empirically-oriented Canadian labour economists worked primarily with U.S. data. Many Canadian faculty members and graduate students knew more about the workings and evaluation of the U.S. welfare and unemployment insurance systems than their Canadian equivalents. At a very modest cost, CERF has effectively and rapidly eliminated this problem, and has created a vibrant, interactive community of academic and government researchers. More information about CERF's activities is provided in **Appendix E**. (see p. 29)

We believe that CERF's formula can be adopted in other areas of inquiry, and applied to a broader range of activities designed to bring together researchers and policy analysts from a wide range of institutions. A successful forum requires a critical mass of researchers in a well-defined but not too specialized area of research. The main activity of the forums would be to organize conferences, with the goals of informing policy makers and policy analysts of research results, providing researchers with a better understanding of policy debates, and more generally offering a setting to facilitate contact among the different sectors of the Canadian Social Statistics Research System. We also see this as an important opportunity for bringing graduate students and beginning researchers into a research milieu. Forums could also provide a convenient and effective vehicle whereby researchers in both the academic and government sector could discuss advantages and limitations of Statistics Canada's surveys for addressing current policy questions, and transmit these insights to Statistics Canada.

We propose that SSHRC conduct peer-reviewed competitions for research forums in quantitative social science. Each forum would be funded for a period of five years, and would be renewable as long as the objectives of the program were met. We envisage the creation of one or two forums per year in the initial five years. For simplicity, these applications could be assessed by the same committee that assesses Research and Training Groups. To qualify as a forum, a group of researchers would form a steering committee, of approximately a dozen members, who would be co-applicants. In addition to the university-based researchers, the appointment of researchers from government and the non-profit sector to the steering committee should be encouraged. Participation and, where possible, financial support from other organizations, including universities, government departments and the non-governmental organizations concerned with public issues, should be viewed positively in assessing applications.

Applications for new forums should include a plan for an inaugural conference in the first year, plus an indication of its direction for the following years. Of course, existing groups with activities of the kind we have described may apply for support. The funding is mainly required for a forum's key activity – organizing conferences – and will include some secretarial and logistic support, travel (both for university faculty and graduate students) and small honoraria for the preparation of conference papers that involve a major departures from a researcher's ongoing research program. To circulate conference papers and maintain communications, funds should cover the maintenance of a web site; and some support should also be

available for publicizing the conferences. Because a successful forum will depend heavily on the largely volunteer efforts of academics in organizing conferences, building and running the organization, some teaching release for Canadian academic members of the steering committee should be available. Travel funding is also required for attendance at meetings of the steering committee.

B. Social Statistics Communications Program

> We recommend that SSHRC establish a Social Statistics Communications Program, specifically aimed at increasing public awareness of quantitative social science research. In co-operation with Research and Training Groups, with Research Forums, and with Senior Fellows, and in consultation with the communication services of universities and of Statistics Canada, this Program should work towards raising the profile of social statistics research findings, and towards furthering debate around the publication of such statistics.

Social statistics are a privileged area for communications in the social sciences, because of the steady stream of quantitative information about social and economic phenomena being published in the media. Various sections of the public, as well as non-governmental organizations, have also come to expect these numbers and the debates that accompany their announcement. Given this favorable context, social statistics should become a priority in the communications strategy of SSHRC. Their Social Statistics Communications Program should be articulated around two streams of information, one coming out of research in academia, the other from the regular publications of Statistics Canada.

In the first place, university researchers produce a number of research findings that are of great interest to the public and to policy-makers, and this will increase as Research and Training Groups come into being. The paradox here is that communications are best managed centrally, or at least from a few key locations where the information can easily be made available to all who could use it, while the research supported by SSHRC's programs is produced in a multitude of research groups and centres, distributed in a number of institutions of higher learning. The challenge, then, is to gather and organize the information, and to prepare researchers for its effective transmission, while seldom having direct control over the way in which communication does take place.

The Communications Program should first monitor the flow of research involving social statistics, using for this purpose connections to Research and Training Groups, to Senior Fellows in Social Statistic and, above all, to Research Forums. Indeed, the Program should be systematically represented on the Boards of the Research Forums. Most of these can be expected to have web pages, and there should be a central web page of these web pages. The media's attention should also be drawn to interesting and relevant research being produced or presented in these places; a bulletin could be published regularly, including for each item the name of reference persons that can be contacted about the findings. Moreover, the Program should have regular relationships with the communications services of universities, encouraging them to raise the profile of social statistics research. This could be achieved both by having universities signal such research to the Program, and by the latter providing support, when needed, to these university services. Finally, the Program could offer training workshops to researchers, for instance at the Congress of Social Sciences, on how to deal with the media.

In the second place, Statistics Canada regularly publishes a variety of social and economic statistics, thus bringing issues to public attention. When the profile of specific issues has thus been raised, SSHRC could take advantage of the moment to make the ideas of academic social scientists more visible to the public and the media. While the purposes and operations of the Communications Program at SSHRC will remain quite distinct from those of the Communications Division at Statistics Canada, some co-operation would be helpful. For instance, given some advance notice about Statistics Canada's publication calendar, the Communications Program could present the media with a current list of experts in any given area of research, and help identify potential contributors to debates on social statistics. Given the decentralization of many aspects of social policy, attention could be paid to identifying experts coming from the various

regions of Canada. The Program could itself organize presentations, and debates among experts, on issues where social statistics can be used to throw light on issues. The public for these events would often be the media, but it could also be public servants, politicians or various organizations. Co-operation and a division of labor with the Humanities and Social Sciences Federation would be useful in this regard.

Finally, the Program could, in co-operation with the Research Forums, organize lecture tours and workshops where research findings based on quantitative social science analysis would be presented to various groups, especially to potential new researchers in social statistics in universities. This activity could be undertaken in co-operation with Statistics Canada, which is already active in this field.

V. Co-ordination, Costs and Timetable

A. Co-ordinating the Canadian Social Statistics Research System

> In order to co-ordinate the implementation of the program in this report, and to monitor progress, we recommend that SSHRC and Statistics Canada agree to a memorandum of understanding specifying their short- and long-term objectives under this initiative. The two organizations should also jointly appoint a co-ordinator for the overall Research System and an advisory committee of senior officials from academic and other organizations.

The initiative in quantitative social research we have described involves an ensemble of interrelated programs. A strong research environment involves a set of mutually-reinforcing institutions, which is why we refer to an emergent "Canadian Social Statistics Research *System*". The ongoing coordination of various components could benefit from an explicit agreement on goals and the formal coordination of the program elements. The actors filling the different roles in the system will need more systematic information about each other than informal networks can provide. There are also some overarching issues that do not fall within the mandate of any one of the components of the system, such as efforts to foster intellectual and methodological perspectives that cut across all research areas and activities. We recommend that this goal-setting and coordination take the form, first, of a memorandum of understanding between SSHRC and Statistics Canada, and, second, of a position for a co-ordinator of the initiative.

We recommend the appointment of a full time co-ordinator, at the director level. The co-ordinator should have extensive experience in research brokerage and management, as well as the intellectual breadth and research experience required to communicate effectively with a range of social scientists. Part of the co-ordinator's mandate should be to seek out and develop partnerships that will facilitate the overall research effort.

The co-ordinator could convene a liaison group, involving representatives from all the components of the system : the chairs and the administrative officers of the peer review committees dealing with the Research and Training Group, fellowships and forums programs; representatives of the Research and Training Groups and Research Forums themselves; co-ordinators of the Summer Workshops and Research Data Centres; a representative of the communications program. The co-ordinator would report to the Chief Statistician and the President of the Social Sciences and Humanities Research Council.

In order to build and maintain strong linkages to the academic and policy research communities, it is also recommended that a senior level advisory committee be formed. Representatives on such a committee should include a number of senior representatives from the academic community (e.g. Vice-presidents for research) and government (e.g. Assistant Deputy Ministers for policy), as well as international researchers or research managers from countries that have successfully implemented similar research initiatives.

B. Costs and Timetable

The proposed research initiative consists of a number of inter-related components. The Joint Working Group did an initial costing of the various components by considering the costs of similar programs. We recognized that a much more extensive costing exercise would be required as the proposals are transformed into detailed activity work plans.

We also recognized that the full program will take a number of years to reach full development, and the assumption has been made that the program would be phased in over a five-year period beginning in fiscal year 1999-2000. The Working Group feels there is a pressing need to proceed quickly on a number of components, and that work could and should begin in this next fiscal year. The highest priority should be given to the start-up, at least on a pilot basis, of several Research Data Centres. It should also be possible to initiate the training program by beginning with a few selected courses, and to launch a few research forums by holding a number of research conferences on priority topic areas. Work should also begin at SSHRC on a communications strategy with respect to social statistics. It is recognized that major competitive programs, in particular the Research and Training Groups and Senior Fellowships, will require a longer review process; they should consequently be started as soon as possible.

Appendix A

Some Priority Areas for Research in Social Statistics

A wide range of policy issues and research areas could benefit from the strengthening of quantitative research which would put to use the available databases. The role of the Joint Working Group, however, is not to "pick winners," but to establish programs with fair selection processes. The brief topic descriptions which follow are intended to suggest, but not restrict, the research that this initiative would stimulate.

1. Child Development

In the past ten years, there has been an increased awareness that the quality of children's experiences during the formative years has long-term effects on their happiness and well-being, their future educational and occupational experiences, and their health status as adults.

A research agenda on early childhood development in Canada could focus initially on the following questions:

- What is the prevalence of Canadian children who are vulnerable to unduly negative life experiences stemming from poverty, family violence, inadequate parenting, or racial and ethnic prejudice?
- To what extent is childhood vulnerability related to family structure, especially single versus two-parent families, and socioeconomic factors, such as family income and parents' education?
- In what ways do the answers to these first two questions depend on the age of the child, and the cohort?
- What are the buffering mechanisms or protective factors associated with healthy child development?

As the policy community, including parents, teachers, administrators, and government policy-makers, attempts to design a new social policy for Canada, it must figure out ways to strengthen families and communities without dramatically increasing government expenditures. Among practitioners, there is a sense that clinical interventions for all children at-risk are too costly and do not adequately meet the needs of all of them. But are interventions targeted for particular groups a better alternative, or would universal programs likely have a stronger impact? If so, what types of programs would be most effective? Research on child development could provide a means of monitoring our progress towards reducing childhood vulnerability and redressing inequalities.

To address these questions, the chief source of data is the National Longitudinal Survey of Children and Youth (NLSCY).

2. Youth in Transition

The transition from high school to post-secondary education and from education into the labour market is problematic for many youth. While the research literature identifies the family background factors mainly associated with poor academic and occupational attainment, considerable research is necessary to understand the pathways to success and critical transitions for youth between the ages of 15 and 25. Some of the principal questions that might guide a program of research include:

- Which groups are particularly prone to leaving school before graduation, and have the most difficulty in making the transition to post-secondary education or to the labor market?
- What are the skills, attitudes and behaviors of youth who achieve successful transitions and those who do not?
- To what extent do graduation rates vary among schools within each province? What school-level factors contribute to successful graduation rates and to high academic achievement?
- Do programs such as co-operative education, mentoring programs, distance education and apprenticeships help students make these transitions?

At least five national and international data sets can be brought to bear on these questions: the National Longitudinal Survey of Children and Youth (NLSCY), the Program of International Student Assessment, the planned Canadian Youth in Transition Survey, the International Literacy data, and the Education and Training Surveys.

3. Families in Flux

Over the last three decades, family life has changed profoundly. The increasing proportion of women in the labour force (especially mothers with young children), as well as the growth of flexible and atypical employment have modified the gender division of labour, both within families and within society, and led to a reorganization of family time. The rise of separation and divorce, the decline of marriage, and the increase in cohabitation have transformed the family trajectories of women, men and children. Researchers have just begun to understand the far-reaching implications of these changes. The questions to be answered include:

- What are the effects of the changing labor environment on the propensity of men and women to both start and maintain conjugal and family relations?
- How are the existing relations among family members (e.g. between conjugal partners, between parents and children, between children and grand-parents) modified and redefined once the family separates?
- What are the consequences of family disruption on fatherhood?
- How are resources shared between partners once a union is dissolved?

To address most of these questions, longitudinal data are required. Such data sets include the 1984 Family Survey, the 1990 and 1995 General Social Surveys on the Family, the National Longitudinal Survey on Children and Youth, and the Survey of Labour and Income Dynamics.

4. Growing Old in Canada

In recent years, researchers, policy makers and program managers have begun to direct attention to the process of aging and to the status of being old in Canada. Major gaps in our knowledge about growing old in Canada include:

- Good descriptions of the uneven retirement process of Canadians from the age of 50. Canadian policies and research are largely based on the misconception that everyone in the labor force enters at about 20 and retires at age 65.
- Knowledge about the oldest old, aged 85 and over. Important questions concern the health of this group, care giving and receiving, living arrangements and income security of people with limited employment-based pensions (including CPP/RRQ).
- The characteristics of people who will likely enter and are currently in residential institutions. Knowledge about these groups is critical to health, housing and economic security planning and policies.

Numerous data sets exist to address these issues, including the National Population Health Survey, Survey of Labour and Income Dynamics, General Social Survey cycles on related topics, CARNET data, the Longitudinal Administrative Data files, the Canadian Study of Health and Aging, the Survey of Ageing and Independence, the Health and Activity Limitation Survey, and the Residential Care Survey.

5. Education, Skills and Literacy

The development of skills and human capital, along with technological advancement (both "soft" and "hard") are seen as the primary forces driving productivity, and hence the standard of living in modern economies. It is difficult to overstate the importance of education, training and skill development for most societies. While this topic has been the focus of major research efforts, rapid changes in the economy and society, and particularly changes in the role and significance of education and skills requires on-going research. In a "knowledge-based" society, the following issues are currently on the policy agenda:

- The performance of Canadian students in a national and international context
- The effect of changes in the education system on access to higher education

- Lifelong learning and its implications
- The role of literacy, *independent* of educational attainment, in labor market success and daily activities
- The adequacy of training in Canadian firms
- The link between human capital and technological change
- Skills shortages and oversupply
- The role of human capital in wealth development.

A number of existing data sources cover these issues, including Adult Education and Training Surveys, International Literacy Surveys, Academic Achievement Tests, Graduate Follow-up Surveys, the planned Workplace and Employee Survey, and traditional surveys such as the Labor Force Survey as well as censuses.

6. The Distribution of Wages and Work

Developed economies today face central policy issues concerning changes in the distribution of wages and of work. In particular, many countries, including Canada, the U.S. and Britain, have seen a substantial increase in wage inequality in the last twenty years. Key research questions include:

- What has caused these trends: increased imports from low-wage countries, and the related phenomenon of outsourcing to those countries, or the introduction of new technologies that eliminate the jobs of less skilled workers?
- Is the collapse of unskilled men's wages a result of the declining influence of unions? Do changes in the quality of our basic education system play a role? Does the relative supply of highly educated workers influence inequality? If more than one of these factors is at work, what is their relative importance and the pattern of their interplay?
- Does a country's institutional structure affect whether it has an unemployment problem (like France and Germany) or a wage inequality problem (the U.S.), or both (Canada)? Does labor market policy responds to research findings, and if so, how?
- What will the role of social support systems be in the future?

Substantial research has been conducted on these topics, but many questions remain. New data sources, such as the Survey of Labour and Income Dynamics and the Workplace and Employee Survey, as well as more traditional sources such as the monthly Labour Force Survey will shed new light on the role of skill-based technological changes and related questions.

7. Social and Community Supports

Over the past decade there has been an increasing recognition of the importance of unpaid work activities, including childcare, household work, eldercare and volunteer work. Of course, women continue to do the majority of unpaid work, despite the increase in their paid labour force participation. The Federal Government's Policy Research Initiative has identified changing time allocation over the life course and within each stage of the life course as underlying many of today's social policy challenges. Care for both children and seniors is also undergoing changes. Aging of the Canadian population coupled with shifts in the responsibility for care, away from institutions and towards individuals and families, are major challenges. The combination of a declining age of retirement with increased life expectancy may also result in time imbalances at older ages. Questions are increasingly being raised about the erosion of community support or "social capital".

Data addressing these issues include: material in various years of the General Social Survey dealing with time use, and with social support and caregiving; the National Survey of Volunteering and Giving; the national Censuses; the National Population Health Survey; the Canadian Study of Health and Aging and the National Longitudinal Survey of Children and Youth.

8. Social Impacts of Science and Technology on Families/Children and on Well Being

Science and technology are dominating forces in this century, perhaps even *the* dominating forces. They are often argued to benefit quality of life and children's futures, and yet we know little about the longer-term social impacts of technological change. It is difficult to make broad statements about whether their effects are positive or, more generally, to describe how they work. Also, there are poorly understood distributional issues. Who benefits and who loses; for example, what is the long-term effect of the gap between children with school- and home-based access to computers and the Internet?

Little is known about the social impacts of science and technology on families and children, and on well-being. Current efforts by Statistics Canada to construct a coherent framework for the systematic development of statistical information for science and technology present opportunities for analytical exploitation of existing data and the production of new data vehicles and linkages.

A number of data sets exist, or will soon be available, to address these issues, such as the Innovation Surveys, the Workplace and Employee Survey, the Graduates Surveys, special surveys on Internet use, research capabilities, and so on. The development of new data is also necessary to extend research areas.

9. Evolving Workplace and Technology Use

The 1990s has seen the intersection of a number of technology-related phenomena that affect the workplace and workers, among which:

- Increased use of information technologies, including rise of the Internet and related communications technologies in almost all industries, accompanied by a concern about whether this has led to the desired productivity gains
- Increased focus on the importance of innovation for firm survival and growth, and for productivity gains
- The effect of technology on downsizing
- Concern that technology may be leading to increasing polarization in society
- A focus on human resources issues such as training, pay practices, work schedules, and new workplace practices implemented to achieve "high performance" workplaces.

Relatively little is known about the adoption and diffusion of technology and innovation in work organizations and its effects on the workplace and workers. While researchers have long been concerned with these issues, few large-scale data sources have existed to document the rate of implementation of technology and innovation, let alone its implications. This has led to the use of often-questionable proxies for technology use, or case studies.

More recently new bodies of data have evolved to assist in our understanding of issues in this area, and to provide new research opportunities. Such surveys include technology surveys, innovation surveys, the Workplace and Employee survey, and a survey of the determinants of firm growth.

10. Welfare, Income and Poverty

Research on welfare and material inequality addresses three main questions. First, what is the distribution among individuals, families, household units and "communities" of income, wealth and the necessities of shelter and food? Second, what are the personal consequences of this inequality on the quality of social life, generally and on specific issues such as health? The third question is what mechanisms reproduce, or alter, inequality over time, both over the lifetime of individuals and between generations? At the aggregate level, this question can be reframed in terms of the evolution of geographical and other forms of communities through time?

Studies of welfare, income and poverty are necessary to understand the effects of a very wide range of policies involving huge expenditures including: the redistributive effects of taxation, the efficacy of social welfare programs, education and health programs, and arrangements for the delivery of social services and health care.

Basic profiles of economic and social inequality have been available for several decades. But as policies and programs change, new research is required to assess their impacts. The new longitudinal surveys provide a hitherto unavailable way to examine closely the process of change over time. The impact on the welfare of Canadians of changes such as the loss of a job or the dissolution of a marital union can be addressed. The short and long-term impact of persisting conditions of severe deprivation on the well-being of children and young people, and on their physical and mental health can be assessed.

Data available to address these issues include the Survey of Consumer Finances, the Survey of Labour and Income Dynamics, the National Longitudinal Survey of Children and Youth, the National Population Health Survey, the Family Expenditures Surveys, and censuses.

Appendix B

Curriculum for a Summer School Program

While it may take several years to reach a steady state, it is useful to think about the curriculum of a more mature summer program, distributed over four weeks. For data analysis, the core of the program should be two courses, each two weeks long. The first would provide a basic introduction to regression, designed for participants with relatively little background, with data analysis examples from cross-sectional surveys such as the General Social Survey. A second two-week course, in linear models, would extend regression techniques to censored, categorical, "count" and duration data. The two courses should combine lectures in the morning with mentored practice sessions in the afternoons. In this and other courses, there should be an emphasis, and specific instruction, on how to write about the results of data analysis.

In addition to the two core data analysis courses, on a revolving basis and according to the demand, each summer program should include one or more intermediate-level courses on topics such as longitudinal analysis, hierarchical models, latent variable models, and categorical data. The emphasis would be on building a corps of researchers equipped with modern data analysis techniques to conduct state-of-the-art research. It will often be appropriate to focus on a single survey, especially for longitudinal analysis, where considerable effort is required to gain familiarity with a data set. There should be a mixture of one- and two-week courses, depending on the topic. Some shorter workshops could address, on a rotating basis, more limited analysis topics such as weighting and estimation issues for complex samples, the analysis of pooled cross-sections, and robust statistical methods.

A third kind of methodological course, definitely worth trying, was proposed to the Working Group by two statisticians (one at Statistics Canada and the other university-based). The teaching groups would focus directly on a data set, chosen on the basis of teachers' and students' interests. Students would work as a team, led by a researcher and including a programmer, a survey methodologist, a specialist on the survey topic and a statistician. The course would revolve around using the chosen data to address a specific research problem, with design, statistical and substantive issues addressed as they arose.

Consideration should be given to staging a workshop, perhaps a week in length, on the philosophy, logic and strategy of data analysis. In recent years, there has been more interest in trying to make explicit, and consequently debatable, the thinking underlying the quantitative analysis of social data. For example, a considerable debate over the relative roles of description and causal interpretation of social phenomena was sparked by the publication of Stanley Lieberson's "Making It Count". Such a course might combine an introduction to the problem of what and how much particular data and data analysis can tell us about social processes, with a discussion of strategies for analysis, focussing particularly on complex survey data sets and on how to do analysis that does justice to a complex topic, but is still feasible.

In addition to the applied statistics courses, each summer session should include at least one course geared to developing analysis of a single survey, such as Survey of Labour and Income Dynamics, the General Social Survey on a topic, the National Population Health Survey, or the National Longitudinal Survey of Children and Youth. These could be open to researchers prepared to make a commitment to conduct a piece of research on the relevant data set and should be led by two or more researchers working with the data set; they would also involve an explicit mentorship arrangement for junior researchers. More senior researchers would take the course mainly to familiarize themselves with a data set and might collaborate in mentoring. According to the topic, the data-familiarization should be paired with a statistical workshop dealing with methods required to analyze the data. Where appropriate there might be follow-up activities, including a strategy for providing ongoing methodological advice, plans for reviews of manuscripts and, if the resources can be found, a later meeting to discuss the research results. Depending on the topic and the extent to which the course emphasized training, a course of this kind could last for one or for two weeks.

Each summer session should include at least one course, perhaps a short seminar, oriented around a substantive research topic. The idea would be to gather researchers to review work in an area and discuss research priorities. This should be coupled with a broad examination of the available data sources relevant to the topic. Hopefully one would bring together a mix of junior and more senior researchers to stimulate

research in the chosen area and the organizers should develop a strategy to encourage ongoing communication among the researchers.

On a revolving basis, the summer program should offer courses on a variety of topics. An attractive idea would be a seminar on policy research with quantitative databases. Researchers from government departments, as well as policy organizations, could be brought in to describe their own work and current policy and research concerns in their areas. A short course on data concepts should also be considered. The idea would be to focus on the way that Statistics Canada, and survey researchers in general, conceptualize and measure the parameters of key social variables. Some obvious candidates include race and ethnicity, labour force participation, work experience and perceptions of the quality of the social environment.

Because there appears to be no Canadian graduate program with advanced training in survey design, consideration should be given to including one in the summer program. Training in data analysis should produce researchers who are aware of the limitations of their data, but is not a substitute for knowledge of the extensive literature on survey design. Ideally, a course would begin with a review of research on questionnaire design and strategies for pre-testing, evaluating and revising new surveys, then turn into a workshop where the participants would write, test and revise a new questionnaire. The second stage of the course would ideally be conducted in co-operation with Statistics Canada or a university-based survey research organization.

Appendix C

Research Data Centres

To facilitate access to confidential micro-data for research purposes, researchers would become "deemed employees" of Statistics Canada. There are numerous legal and organizational issues that must be addressed under this scenario, and this appendix outlines one approach. One issue is physical proximity to the data. Access is difficult for many if the data are available only in Ottawa. Hence, it is proposed that this approach be developed within a framework of Research Data Centres.

1. What is a Research Data Centre?

A Research Data Centre would be a physical location that would have a secure environment capable of protecting confidential micro-data files, and would be an extension of Statistics Canada. It would have an affiliated research program administered by SSHRC and Statistics Canada. Researchers sworn in under the Statistics Act would have access to confidential micro-data files maintained at the centres, thus allowing research to proceed. Researchers would be held accountable for the protection of confidentiality in exactly the same way as Statistics Canada employees are currently held accountable.

2. The Major Issues Associated with Developing a Research Data Centre Program

The protection of confidentiality. This is of central importance, since confidentiality is the cornerstone of the statistical system. Any access to data by researchers who are not regular employees of Statistics Canada must be done in accordance with the Statistics Act.

Ensuring that the work and the researchers fall under the Statistics Act. In order to ensure confidentiality, the researchers must be sworn in under the Statistics Act, and be subject to the conditions of the Act in the same manner as a regular Statistics Canada employee. The Act lays out the conditions under which people can be given access to confidential data, and hence these conditions must be incorporated into any program of the Research Data Centres.

Ensuring that the Centres will succeed in their objective. The centres must be structured in such a way as to attract top quality researchers and their students to ensure that the goals outlined earlier in the report are met.

Ensuring an appropriate organizational structure. There are issues regarding the manner in which such data centres are organized and managed and the composition of the approval and evaluation committee for the research proposals. It is proposed that SSHRC and Statistics Canada jointly administer the affiliated research program, and that Statistics Canada manages the facility, as it is an extension of that agency. Operationally, the selection and vetting for the research program would be conducted by committees consisting of prominent researchers, in compliance with the requirements of the Statistics Act and confidentiality.

Ensuring appropriate physical safeguards for the data. To both meet confidentiality requirements and to openly demonstrate that they are being met, adequate physical security obviously must be in place.

These issues are addressed in the following outline.

3. The Goals of the Research Data Centre Program

The program would have two goals:

- To promote quantitative research by academics, researchers in government agencies, research institutes or elsewhere in the public sector using Statistics Canada micro-data files, particularly household micro-data files.
- To improve the statistical programs of Statistics Canada through feedback from researchers using the micro-data, and through the research and papers produced by the researchers.

4. Elements of the Research Data Centre Program

Following are the essential elements of a research data centre program:

- A secure, but user-friendly computer environment in which confidential micro-data could be stored would be required. Security would be at the standard maintained by Statistics Canada.
- A selection, approval and vetting process for the research based on confidential micro-data. A committee consisting of senior researchers (from academia and elsewhere) in a number of fields, as well as SSHRC and Statistics Canada officials would be created to administer the selection and vetting process. By and large, senior academics would determine which projects would be approved. There would be one peer review panel for each major area or research, such as economics, sociology, health, education, and statistical methodology[1]. Academics or researchers at research institutes, government agencies, or other research organizations could submit proposals.
- There would be at least one Statistics Canada employee on-site at the Data Centre, to manage the site, provide support and to oversee confidentiality issues.
- Only persons sworn in under the Statistics Act through the Research Data Centre Program would have access to data at the Centres. Researchers would sign a contract that dealt with their obligations under the Act[2].
- The Centres would have to be self-financing, with funding coming from the institutions running the Centre, granting councils such as SSHRC, or through the researchers.

5. Operating the Centres under the Statistics Act

Researchers with access to confidential data must be sworn in under the Statistics Act. Aside from regular Statistics Canada employees, the Statistics Act restricts access to confidential data to people who are "retained under contract to perform special services for the Minister (i.e. Statistics Canada)." These people become deemed employees of Statistics Canada. That agency retains researchers to conduct work when it does not itself have the resources to do all the work required. Statistics Canada clearly does not have sufficient resources to produce the analytical work needed to exploit the many new (and older) data sets. Such exploitation is needed to provide analytical insights of value in public policy development and debate, and in the promotion of basic research. Much of the capacity to take advantage of the rich data resources resides in the academic community and in other research organizations (in other federal agencies, think tanks, provincial governments, etc.).

In this environment, and under the current Act, in order to be sworn in under the Act the research done would be similar to that which Statistics Canada itself would normally conduct (if it had the resources). While this may sound restrictive, in fact Statistics Canada carries out a wide range of research; consequently, this is not likely to be an issue. In the selection process, the merits of the proposals would be paramount. This selection and review committee would oversee the review process after the completion of the project. Statistics Canada would obviously have some input to the process, since under the Statistics Act the work constitutes special services for the Minister.

Seeking approval from respondents to share the data with data centre users. There is an alternative means by which confidential data can be shared with users who are not covered by the Statistics Act. The Act includes a provision for the sharing of data with users. If the respondent's permission is sought at the time of data collection, confidential micro-data may be shared with selected users. Data sharing would be possible with SSHRC or an "Institute" created by that agency. Respondents would be informed about the nature of the institute (or SSHRC), and permission sought to share the data. Based on past experience, it

---

[1] Preference might be given to empirical research that has policy relevance. To allow the system to be responsive to rapidly emerging policy questions, a two-tier approval process should be considered. One tier would be designed to render quick decisions for short-term feasibility studies with only minimal funding, and another directed at longer-term projects with more substantial resource implications.

[2] In order to promote the training of a new generation of researchers with expertise in the use of these data sets, and in order to facilitate the research itself, this should include research assistants, especially graduate students, employed by the academic researcher or under his/her supervision in dissertation research.

is likely that most respondents would agree. Under such a scenario, Statistics Canada would then provide the raw micro-data to the Institute (or SSHRC), which would of course agree under a contract arrangement to maintain confidentiality.

This approach has the advantage of being straightforward. The selection and vetting process would then be entirely in the hands of the Institute; Statistics Canada would play little role other than setting up the conditions to ensure data confidentiality. The shortcoming of this approach is that it can only be applied to data collected in the future, and to some data sets. The data sets that have been created to date could not fall under such an arrangement. For that reason, this approach should be seriously considered for the future, but will not solve the current issues. Thus, there is the possibility of forming an "Institute" with which data sharing could take place. The propensity of respondents to share the confidential data with such an institute (or SSHRC) for research purposes could be tested. While this approach would have little benefit in the short run, in the longer run it is a potentially excellent solution to sharing selective data sets.

Handling the research paper upon completion of the project. Researchers may want to think in terms of two papers following completion of the project. A research report consisting of the quantitative analysis and an interpretation of the findings would be deposited with the Selection and Review Committee of the Data Research Centres. This product would become part of a research paper series[3], and would be reviewed in a normal academic peer review manner. The Committee (or it's named designate) would run this process. The product would also undergo an "institutional" review by Statistics Canada. This is simply to determine if the work falls within the mandate of the agency for the purposes of the research paper series. Work conducted for the research paper series could be (and often should be) policy relevant but could not contain direct policy recommendations. Hence, the institutional vetting by Statistics Canada is conducted primarily to ensure that there are no direct policy recommendations in the paper[4]. After the academic and institutional reviews, authors will make the necessary changes.

The researchers would, of course, be free to publish the research paper, or an extended version of the research paper with policy comment and other additions they see fit, in any academic journal or any other venue. In short, beyond the initial quantitatively-oriented paper that is submitted to the research paper series, the researcher would be free to produce any other version of the paper, and submit it for publication in any forum. In the event that the review of the paper by the Selection/Evaluation Committee and Statistics Canada leads to a decision not to publish the paper in their research paper series, the author will be able to submit if for publication elsewhere[5].

---

[3] The research paper series developed for the publication of the initial quantitative analysis would likely be best maintained in one central location. Managing the research paper series centrally ensures that there is some consistency in the way the papers are handled. There may be resource considerations for Statistics Canada and the Committee. Vetting the research papers to ensure that the work is properly conducted may at times be time consuming. Such vetting will be largely done through the academic refereeing process. At times, however, work may be required to validate the use of the data. It is important that the person responsible for maintaining the series understands research and the research world, and has the resources necessary to prevent a backlog of papers.

[4] Direct policy recommendations or comments refer to direct evaluation, criticism or advocacy of existing or proposed government programs. This does not exclude research on policy relevant topics, which is in fact encouraged. These restrictions are placed on Statistics Canada output (and the joint research paper series) to protect its neutrality and objectivity.

[5] Normally the copyright on work that has been done under contract as a special service for the Minister (as required by the Act) would be vested with Statistics Canada. However, in the contract struck before the work begins, it is proposed that Statistics Canada agree to vest the copyright with the researcher. Statistics Canada would retain the right to vet all publications stemming from the project for confidentiality and data reliability. In the same contract Statistics Canada would retain the right to reproduce the paper if it chose to do so.

6. Where the Centres Might Exist

The centres may be affiliated with a university, research institute, or research network. They could also be located at Statistics Canada regional offices, and for reasons outlined later, it is proposed that regional offices be employed initially. Obviously a physical location with a secure environment is needed. If located outside of Statistics Canada, competitions could be held to determine where such centres might be placed. One would want to start with a very small number, as there would be a substantial resource impact on Statistics Canada through the servicing of such centres. It seems likely that having such a centre would allow the institute serving as a home base to develop a very strong empirical research capacity in potentially a number of disciplines. Very qualified empirically oriented researchers would be attracted to an institute with a Data Centre, allowing the university or institute to develop a strong program.[6]

7. Starting Small

This is an ambitious program. A pilot approach may be appropriate. It is proposed that the data sets made available to the centres initially include the new longitudinal surveys: the Survey of Labour and Income Dynamics (and the Labour Market Activity Survey, its predecessor), the National Longitudinal Survey of Children and Youth, the National Population Health Survey, the Workplace and Employee Survey, and related social statistics data sources required to conduct the research. It is here that the need for direct access to the micro-data is the greatest. It is also proposed that regional offices be the initial sites for the pilot project. This would allow the development of the centres to proceed incrementally. Nonetheless, access to the confidential micro-data would be improved tremendously, and some fairness regarding the geographical location of the centres would be introduced, as regional offices are located across the country. If housed in Statistics Canada, the initial centres may be less costly to run, as some of the infrastructure already exists. After the pilot project has been in place for some time (perhaps after two years), it is proposed that the program be reviewed, and subject to this review, a competition for Research Data Centres at Universities or other research organizations be held.

8. The Funding

Similar centres operated by the U.S. Census Bureau and the U.S. National Science Foundation have an annual budget of U.S. $250,000 per location. There are at least two alternative funding approaches. The first is similar to the way in which the Data Liberation Initiative (DLI) is financed. Under this scenario, the Centres would be entirely block funded. Such block funding would be provided by SSHRC, the university or institute at which the centre is located (after the pilot phase), and other organizations whose researchers use the centre, such as government departments, research institutes, etc. Since it anticipated that primarily academic researchers would use the centres, the majority of the block fund would be SSHRC-based. There would be no direct cost to the researchers using the centres. The selection process run by the selection and review committee would regulate access to the centres. A variant of this approach would be to have much of the annual cost covered by block funding, and the remainder covered by fees paid be researchers using the centre. University-based researchers whose projects were approved by the committee would receive a SSHRC grant (distributed through the review committee) to cover this cost. Researchers from other organizations would pay similar fees. It is proposed that one of these funding approaches be implemented.

[6] Certainly the NBER (National Bureau of Economic Research) in the U.S., one of the first institutes to have a Census Bureau Data Centre, has developed a very strong empirically based research program, presumably in part due to the superior data access available to researchers affiliated with the NBER.

9. Conclusion

This proposal urges the consideration of Research Data Centres as one solution to the issue of access to micro-data for researchers. Initially this would involve access to household survey data, particularly the new longitudinal household surveys. Business (establishment or company) surveys present unique confidentiality issues that may prevent them from being included in this endeavor, at least initially. The Centres could be supplemented by a remote access capability developed within Statistics Canada. The latter would be useful for researchers who are not geographically capable of using a Data Centre, or who for other reasons cannot or wish not to use the Data Centres. Smaller projects, for example, may be better dealt with through remote access.

Appendix D

The Development of a Remote Access Capability at Statistics Canada

A Remote Access Capability should provide researchers with tools allowing them to specify statistical procedures and to have these procedures applied to confidential data by Statistics Canada staff. The results would be screened by Statistics Canada for confidentiality and returned to the researchers.

Among the tools available to the researcher would be the public use micro-data file (where available) and a permuted file which closely mirrors the confidential file in structure and detail, but with only a limited amount of real data. The researchers would specify their program or desired run as an SPSS/SAS job, using the permuted file, and then submit it to Statistics Canada via the Internet. The permuted files should contain sufficient detail to provide a basis for the testing and debugging of programs. Statistics Canada would not play a role in such debugging; program errors would be returned to the researcher. The main task of Statistics Canada employees would be to run the jobs against the master file and vet the output for confidentiality. Programs that led to confidentiality problems would have to be modified by the researcher

The main advantages of remote access are its availability to researchers regardless of their physical location, and the simplification of the interface between them and Statistics Canada. Its disadvantages are its costs (when compared to the direct use of publicly available data files) and the possible delays incurred as the results are vetted. An informal working group has been struck within Statistics Canada to develop an implementation strategy that would minimize these disadvantages. For the moment, it goes as follows: the files would be produced by the subject area responsible for the survey, with the assistance of methodology staff skilled in the generation of such files. They would be distributed to researchers via the Data Liberation Initiative (DLI)/ FTP site; DLI contacts at universities could help to facilitate access to the files as they do with the current DLI files. Statistics Canada would establish a central group, building on the present DLI team, to run the programs and screen the outputs for confidentiality. The results would be returned to the researchers using FTP.

The main challenges about this approach have to do with the range of permuted files that can be produced and made available, and with the effectiveness with which this can be done, with the effort required to screen the output for confidentiality, and with the turn-around time. Many of these questions will be answered only after a suitable test of the strategy. The latter has been implemented for the National Population Health Survey (NPHS) and it is anticipated that there will be four files available for testing in a remote access environment next year. This should provide researchers and Statistics Canada with enough information to decide how remote access can be used, in conjunction with other approaches, to improve access to data.

Appendix E

A Model for Research Forums in Social Statistics: the Canadian Employment Research Forum

CERF is a non-profit corporation, whose primary goal is to increase productive interaction between researchers studying employment issues in the academic and government sectors, as well as policymakers themselves, and to increase both the volume and quality of policy-relevant research in this area. CERF is directed by a rotating volunteer board, at least one third of the directors come from the academic community and from the government service. CERF receives core funding from Human Resources Development Canada, but it also raises considerable additional funds for its conferences from granting agencies, foundations and stake-holding government departments.

In the last eight years, CERF has organized a series of conferences (close to twenty, so far) on a wide variety of research topics. As a result, CERF has dramatically improved the access of policy makers to the technical expertise of academic researchers, at very low cost. CERF also sparked a dramatic increase in research using Canadian data. This was no accident: CERF conference organizers encouraged Canadian research, though they recognized the benefit of comparisons with the U.S. and other nations. Through contacts in government, CERF also facilitated access to confidential data. A list of conferences organized by CERF follows.

- Founding CERF Conference. Kingston, Ontario, May 31, 1991.
- Policy Research in Training, Unemployment and Income Support, and Immigration. Aylmer, Quebec, March 6-7, 1992.
- Labour Markets in the Last Two Recessions: A Comparative Perspective (with Statistics Canada). Ottawa, Ontario, March 5, 1993.
- Income Support Programs and Policies (with Health and Welfare Canada). Ottawa, Ontario, September 24, 1993.
- Immigration and the Labour Market. Hull, Quebec, March 11, 1994.
- Labour Markets and Income Support: Focus on British Columbia (with B.C. Ministries of Social Services and Skills Training). Vancouver, B.C., March 25, 1994.
- Workshops in conjunction with the Learned Societies Annual Conference. Calgary, Alberta, June 13, 1994.
- Youth Labour Market Adjustment (with BC Ministry of Education and UBC), Vancouver, B.C., June 25, 1994.
- International Conference on the Evaluation of Unemployment Insurance. Ottawa, Ontario, October 14-15, 1994.
- Policy Responses to Displaced Workers (with the Université du Québec, Montréal and Ekos Research Associates). Montreal, Quebec, December 2-3, 1994.
- Retooling the Employed Workforce: Focus on New Brunswick (with HRDC, NB Region and UNB). Fredericton, N.B., March 31 - April 1, 1995.
- Environmental Policies and Labour Markets (with HRDC Innovations). Ottawa, Ontario, Sept. 15-16, 1995.
- The Canada - U.S. Unemployment Rate Gap (with Centre for the Study of Living Standards). Ottawa, Ontario, Feb. 9-10, 1996.
- Changes in Working Time in Canada and the United States (with Upjohn Institute, Statistics Canada, HRDC). Ottawa, Ontario, June 13-15, 1996.
- Social Experiments, Evaluation and Employment and Social Policy. 1996.
- Immigration, the Labour Market, and the Economy. Richmond, BC, October 17-18, 1997.
- Participation Rates and Employment Ratios. Ottawa, Ontario, December 17, 1997.
  Labour Market Transitions and Income Dynamics (co-sponsored by Statistics Canada and HRDC). In conjunction with the CEA meetings in Ottawa, Ontario, May 28-29, 1998.
- Self Employment Conference (with the Canadian International Labour Network and the OECD). Burlington, Ontario, September 24-26, 1998.

More information on CERF is available at its website: http://cerf.mcmaster.ca

Appendix F

Membership of the Joint Working Group

Paul Bernard, Chair
Département de sociologie
Université de Montréal

Peter Kuhn
Department of Economics
McMaster University

Douglas A. Norris, Director
Housing, Family and Social Statistics Division
Statistics Canada

Garnett Picot, Director
Business and Labour Market Analysis
Statistics Canada

Martin Wilk, Former Chief Statistician
Statistics Canada

Betty Havens
Community Health Sciences
University of Manitoba

Céline Le Bourdais
INRS – Urbanisation

Michael Ornstein
Institute for Social Research
York University

J. Douglas Willms
Atlantic Centre for Policy Research in Education
University of New Brunswick

Assisted by:

Hélène Régnier, Policy Analyst
SSHRC

Rapport final du

Groupe de travail mixte du

Conseil de recherches en sciences humaines du Canada

et de

Statistique Canada

sur l'avancement de la recherche utilisant les statistiques sociales

Décembre 1998

Résumé

L'économie et la société canadiennes traversent une période de changements rapides et difficiles. Il nous faut disposer, au moment opportun, d'une analyse objective de la conjoncture économique et sociale afin de comprendre la transformation qui s'opère, de jeter les bases d'un débat vaste et informé sur les politiques gouvernementales, et d'établir les fondements d'une méthode judicieuse d'élaboration des politiques. Cette nécessité se fait d'autant plus sentir que la politique sociale du Canada n'a pas évolué au même rythme que sa politique économique, qui s'est métamorphosée au cours des deux dernières décennies. Tous les paliers de gouvernement ont admis l'importance de redéfinir la politique sociale de façon qu'elle réponde aux besoins de l'ensemble des Canadiens, et qu'elle amène le développement de collectivités plus favorables à l'épanouissement de l'individu et plus viables sur le plan économique.

D'un certain point de vue, le Canada est bien équipé pour atteindre cet objectif. Nous avons aujourd'hui accès à un bon nombre d'excellentes enquêtes sociales, exécutées fort à propos, qui se rapportent à une diversité de sujets. Les progrès de la statistique mathématique et de la méthodologie de la recherche ont entraîné la création d'enquêtes longitudinales, qui permettent de suivre un échantillon de personnes sur de longues périodes. Ces nouveaux outils de recherche fournissent de l'information sur la dynamique de la pauvreté, l'efficacité des programmes de formation, les conséquences de la perte d'emploi et les effets des expériences de l'enfance, de même que sur plusieurs autres questions dont l'étude favorise une redéfinition éclairée de la politique sociale. Collectivement, ils constituent la base d'un système bien intégré de « statistiques sociales », terme sous lequel nous désignons tout un ensemble d'information décrivant un large éventail d'activités humaines ainsi que les facteurs sociaux, économiques, scolaires et culturels qui ont une incidence sur la vie quotidienne.

Le Canada ne possède cependant qu'une infime partie des moyens nécessaires pour effectuer des recherches en matière de politique sociale, évaluer ses programmes sociaux ou surveiller l'efficacité des mesures visant des objectifs de société. Le gouvernement fédéral a reconnu récemment qu'il lui fallait renforcer sa capacité de recherche; il a mis sur pied, à cette fin, le Projet de recherche sur les politiques, dont les responsables ont pour mandat de recommander un programme de recherche interministériel et d'en superviser l'exécution. Les gouvernements provinciaux, le secteur privé et les organismes non gouvernementaux sont arrivés à la même constatation, et ont cherché à revitaliser le milieu de la recherche sur les politiques.

Nous avons au moins trois obstacles de taille à surmonter pour développer notre capacité de recherche en statistiques sociales. Le premier de ces obstacles, c'est le manque de chercheurs ayant la formation adéquate. L'entreprise de restructuration et de compression des effectifs, lancée il y a plus de dix ans par le gouvernement fédéral et les gouvernements provinciaux, a réduit considérablement le nombre de chercheurs au sein de la fonction publique. Ce phénomène s'est produit alors que déclinait, dans les départements de sciences sociales de nos universités, la formation relative aux méthodes de recherche faisant appel aux statistiques (à l'exception des sciences économiques). Le deuxième de ces obstacles, c'est la difficulté que présente l'accès aux données. Paradoxalement, les détails fournis par les nouveaux ensembles de données, et qui leur confèrent une grande valeur, empêchent aussi que ces données puissent être rendues publiques, parce qu'elles pourraient permettre à des utilisateurs d'identifier certaines personnes. La *Loi sur la statistique* établit des critères stricts afin de préserver la confidentialité, et Statistique Canada fait forcément tout le nécessaire pour assurer le respect de la *Loi*. Le troisième obstacle, c'est qu'il n'existe que très peu de liens entre les spécialistes des sciences sociales et les utilisateurs éventuels des connaissances qu'ils produisent. Bien que l'on réclame avec insistance des statistiques sociales sur l'éducation, l'emploi, la santé, l'alphabétisation et d'autres questions pertinentes, un bon nombre des découvertes importantes des chercheurs en sciences sociales n'ont pas été communiquées de façon appropriée au milieu où s'élaborent les politiques, ni à la population en général par les médias de communication.

Le mandat de notre groupe de travail, dont les membres ont été nommés conjointement par Ivan Fellegi, statisticien en chef, et Marc Renaud, président du Conseil de recherches en sciences humaines (CRSH), était de formuler des propositions visant à encourager la recherche quantitative sur des questions sociales et économiques d'envergure, au moyen de données à grande échelle. Le présent rapport recommande le financement des trois composantes d'un **Système d'analyse des statistiques sociales**, chacune de ces composantes ayant pour objet de vaincre un des obstacles définis ci-dessus. Le nouveau système

viendrait compléter notre système intégré de données statistiques relatives aux questions sociales, dont l'excellence est déjà mondialement reconnue.

La première composante vise à accroître le nombre des spécialistes qui font de la recherche quantitative sur des questions sociales et économiques. Elle réunit trois propositions distinctes : des **groupes de recherche et de formation**, qui rassembleraient des chercheurs de disciplines et d'établissements différents en vue d'exécuter des travaux de recherche quantitative et d'assurer une formation dans des domaines prioritaires; un **programme de formation**, incluant un programme d'été, qui fournirait des connaissances spécialisées sur les méthodes statistiques de pointe, en complément des programmes de stages et des programmes d'études des deuxième et troisième cycles, ainsi qu'un appui aux bibliothécaires responsables des données, afin d'accroître l'utilisation des statistiques sociales dans la formation des étudiants non diplômés; enfin, un **programme de bourses**, qui comprendrait des bourses de maîtrise, de doctorat et d'études postdoctorales destinées à de jeunes chercheurs faisant carrière en statistiques sociales, ainsi que des bourses de recherche de niveau supérieur qui permettraient à certains des principaux spécialistes des sciences sociales au Canada de consacrer une plus grande partie de leur temps à la recherche, et de jouer un rôle de premier plan dans la formation de la prochaine génération de chercheurs.

La deuxième composante réunit deux propositions : d'abord des **centres d'accès aux données de recherche** et des ressources pour la consultation à distance, qui donneraient accès à des microdonnées détaillées tout en maintenant les règles strictes que Statistique Canada est tenu d'imposer, en vertu de la *Loi sur la statistique*, afin de préserver la confidentialité; ensuite une aide visant l'enrichissement et l'expansion de l'**Initiative de démocratisation des données**.

La troisième composante prévoit la création d'un **programme de communication en statistiques sociales,** qui mettrait en oeuvre une stratégie de communication visant à informer le public et à éveiller son intérêt pour la recherche quantitative en sciences sociales. Élément clé de ce programme, les **Forums de recherche** apporteraient un soutien aux réseaux de recherche, fourniraient une tribune pour la présentation des résultats et amélioreraient la communication entre les chercheurs, les membres du milieu où s'élaborent les politiques et les médias.

Le rapport recommande aussi que Statistique Canada et le Conseil de recherches en sciences humaines négocient un protocole d'entente définissant les buts et le mode d'organisation du Système d'analyse des statistiques sociales qui fait l'objet de notre rapport. Est également recommandée une structure de coordination qui permettrait de tirer parti des occasions de consolider les réseaux et d'accroître la synergie chez les chercheurs.

Table des matières

Page

I. Introduction 1

II. Accroître la capacité de recherche 3

A. Les Groupes de recherche et de formation 4
B. Améliorer les compétences : le Programme de formation et le Programme d'été 6
C. Les programmes de bourses 8
Bourses de maîtrise, de doctorat et d'études postdoctorales 8
Bourses de recherche de niveau supérieur en statistiques sociales 8

III. Faciliter l'accès aux données 9

A. Les Centres d'accès aux données de recherche 10
Créer un service de consultation des données à distance 12
Obtenir des répondants l'autorisation de communiquer les données aux utilisateurs des Centres d'accès 12
B. L'archivage et la documentation des données 13
C. L'initiative de démocratisation des données 14

IV. Communiquer les résultats de la recherche 15

A. Les Forums de recherche 16
B. Le Programme de communication en statistiques sociales 17

V. Coordination, coûts et plan d'exécution 19

A. La coordination du Système canadien d'analyse des statistiques sociales 19
B. Coûts et plan d'exécution 20

Annexe A – Certains domaines prioritaires pour la recherche en statistiques sociales 21

Annexe B – Aperçu d'un programme de cours d'été 27

Annexe C – Centre d'accès aux données de recherche 29

Annexe D – Développement d'une capacité d'accès à distance à Statistique Canada 35

Annexe E – Modèle de forums de recherche en statistiques sociales : le Forum canadien de recherche sur la situation d'emploi (FCRSE) 36

Annexe F – Composition du Groupe de travail mixte 38

I. Introduction

Depuis deux décennies, le Canada connaît des changements radicaux sur plusieurs plans : social, économique et technologique. Notre politique économique s'est adaptée à ces changements, surtout en réponse aux exigences des marchés mondiaux ainsi qu'à la nécessité de lutter contre l'inflation et de réduire la dette publique. Cette évolution de la politique économique s'est accompagnée d'une transformation du marché du travail et de la structure familiale. Beaucoup de Canadiens ont l'impression, en fait, que nos problèmes les plus importants ne sont pas de nature économique, mais sociale. Les gouvernements de tous les paliers ont d'ailleurs reconnu qu'il faut redéfinir notre politique sociale de façon qu'elle corresponde mieux à notre politique économique d'aujourd'hui.

Nous avons besoin à cette fin d'effectuer des recherches sur un large éventail de questions sociales, économiques, scolaires et culturelles. Le renouvellement de la politique sociale doit en effet s'appuyer sur une compréhension fondamentale du cours de l'existence et des relations complexes qui existent entre les facteurs de différents niveaux, comme la famille, le voisinage et la collectivité. La réalisation de tels travaux exige un système bien intégré d'enquêtes statistiques sur des questions sociales, et la capacité d'analyser les données.

En se servant tout d'abord des données tirées du recensement et d'enquêtes sur le commerce, les finances, les prix, la population active et divers autres sujets, Statistique Canada a mis sur pied un système statistique national. Par le passé, son objectif a été de fournir des indicateurs agrégés de la situation économique et sociale, y compris des statistiques descriptives sur la croissance économique, la santé, l'éducation, la justice, la productivité et la population active. Au cours des années 70, deux changements majeurs ont déterminé l'évolution de l'organisme. Premièrement, la recherche en sciences sociales a montré que la compréhension de nombreux phénomènes sociaux, comme les activités criminelles et la victimisation, ou les effets de la pauvreté, demande des enquêtes distinctes et ciblées. Deuxièmement, le milieu où s'élaborent les politiques a reconnu l'importance de travaux de recherche qui peuvent nous aider à comprendre comment certains événements qui surviennent au cours de l'existence, associés aux habitudes des personnes en cause et à leur mode de vie, influencent la situation sociale de ces personnes à long terme. La recherche a commencé à mettre en lumière à quel point le *contexte social* – la famille, le voisinage, l'école, les organismes au sein desquels les gens vivent et travaillent – oriente les actions des individus et les soumet à certaines contraintes. Cependant, les données descriptives fournies par les enquêtes transversales ne permettaient pas d'observer les changements qui se produisent dans la situation sociale, ni de comprendre les mécanismes grâce auxquels on parvient à une situation désirable. Il fallait à cette fin des enquêtes longitudinales, qui consistent à réunir des données auprès d'un même échantillon de répondants à deux occasions au moins, davantage si possible.

Statistique Canada a relevé ce défi en instituant une nouvelle série d'enquêtes sociales, pendant les années 80 et au début des années 90. Il s'agit notamment de l'Enquête sociale générale, des enquêtes de suivi menées auprès des diplômés, de l'Enquête sur la dynamique du travail et du revenu, de l'Enquête longitudinale nationale sur les enfants et les jeunes, de l'Enquête nationale sur la santé de la population et de l'Enquête auprès des travailleurs déplacés. D'autres enquêtes d'envergure doivent s'ajouter au cours des quelques prochaines années, dont une enquête longitudinale sur le lieu de travail et les employés, une enquête longitudinale auprès des jeunes en transition et une enquête sur la sécurité financière. Par ailleurs, Statistique Canada a fait porter davantage d'efforts sur la collecte de données administratives et de données fiscales, et a renforcé sa capacité d'établir des liens entre des données de sources diverses. Des chercheurs canadiens ont joué en outre un rôle de premier plan dans l'élaboration et la mise en oeuvre d'enquêtes internationales, comme l'Enquête internationale sur l'alphabétisation des adultes et la Troisième étude internationale de mathématiques et des sciences. Plusieurs autres enquêtes sont actuellement au stade de l'élaboration, et l'on s'efforce d'intégrer les données recueillies de façon qu'elles reflètent plus fidèlement la vie des Canadiens. Toutes ces enquêtes ont donné une

nouvelle orientation au système statistique national, et ont contribué à faire du Canada un leader mondial pour la réunion de données statistiques sur des questions sociales.

En dépit de ces réalisations importantes, le Canada se montre relativement peu en mesure d'analyser les données tirées des enquêtes. L'obstacle le plus difficile à surmonter, c'est qu'il y a simplement trop peu de chercheurs qui font de la recherche quantitative. Ce problème est particulièrement sérieux dans les domaines qui exigent le recours à des méthodes statistiques de pointe. Au sein des organismes gouvernementaux, la capacité d'effectuer des recherches a baissé au cours des deux dernières décennies, en raison principalement de l'entreprise de restructuration et de compression des effectifs. En outre, bien que ces organismes soient conscients de la nécessité de mettre à profit les analyses de statistiques sociales pour éclairer les débats sur les politiques gouvernementales, ils n'entretiennent pas de liens assez solides avec les universités, où la majorité des spécialistes des sciences sociales exécutent leurs recherches, et où se forment les nouveaux chercheurs. Au sein des universités, la formation relative aux méthodes de recherche utilisant les statistiques a aussi connu un déclin considérable ces vingt dernières années. Il se trouve maintenant très peu de professeurs qui enseignent les méthodes statistiques de pointe et, dans la plupart des disciplines des sciences sociales, les programmes de maîtrise et de doctorat ne comportent plus obligatoirement une formation officielle à l'utilisation des statistiques ou des méthodes de recherche quantitative. L'économique fait notablement exception à la règle, mais dans cette discipline la capacité de recherche est érodée, depuis un bon moment, par le départ de chercheurs qualifiés pour les États-Unis.

Ce problème est d'autant plus grave que les données produites par les nouvelles enquêtes sont assez complexes : ces enquêtes sont généralement de type longitudinal, et ont une structure multiniveau; elles portent par exemple sur des étudiants ou des travailleurs, nichés les uns dans les écoles, les autres dans les entreprises. Il faut donc, afin d'en analyser les résultats, des techniques de calcul et des techniques statistiques très poussées. La statistique mathématique a fait d'énormes progrès dans ce domaine, mais nous comptons au Canada très peu de spécialistes des sciences sociales qui ont reçu la formation nécessaire pour en tirer parti.

La difficulté d'avoir accès aux données constitue un autre obstacle important. On a fait déjà beaucoup de travaux utiles à partir des fichiers de données à usage public que Statistique Canada a mis à la disposition des chercheurs, tant des universités que d'autres établissements, surtout après le lancement de l'Initiative de démocratisation des données par le milieu universitaire, avec l'appui de Statistique Canada, de la Fédération canadienne des sciences humaines et sociales et du CRSH. La cause du problème que nous connaissons aujourd'hui, c'est la nécessité de rendre accessibles aux chercheurs les microdonnées détaillées – les réponses exactes aux questions posées dans le cadre des enquêtes –, sans mettre en danger la confidentialité que l'on a promise aux répondants. Les techniques statistiques très poussées qui conviennent à l'analyse des données longitudinales à niveaux multiples ne s'appliquent pas aux données d'ensemble; la consultation des microdonnées est essentielle. En outre, même de simples problèmes de description exigent dans bien des cas le recours aux microdonnées. L'ironie veut que la difficulté d'accès vienne en partie de ce que les microdonnées longitudinales, en fournissant des détails concernant les individus, pourraient permettre au chercheur d'identifier certaines personnes. Cela contreviendrait aux dispositions de la *Loi sur la statistique* qui exigent de préserver la confidentialité des réponses.

Le troisième obstacle au développement de notre capacité de recherche en sciences sociales se rapporte à la communication. Beaucoup de spécialistes des sciences sociales se consacrent par nécessité à l'étude de questions rigoureusement circonscrites, qui fera progresser les connaissances dans leur discipline. L'avancement d'une carrière universitaire repose en grande partie sur des publications dans des revues savantes, et il existe généralement assez peu de mesures incitatives ou de ressources pour encourager la rédaction d'articles de vulgarisation, qui communiqueraient les résultats de recherches à un vaste public. En outre, comme la poursuite d'une carrière universitaire les oblige à exceller dans la recherche et dans l'enseignement, les chercheurs n'ont guère le loisir de nouer des liens solides avec le milieu de l'élaboration des

politiques, et dans bien des cas la durée des travaux de recherche universitaire est trop longue pour répondre aux besoins immédiats des décideurs. Par conséquent, le transfert des connaissances depuis les chercheurs jusqu'aux responsables de l'élaboration et de la mise en oeuvre des politiques n'est pas aussi rapide ni aussi efficace qu'il pourrait l'être. Peut-être pour les mêmes raisons, les spécialistes des sciences sociales n'ont pas vraiment réussi à tirer parti des médias d'information pour diffuser les résultats de leurs travaux.

Cependant, afin de renforcer à long terme notre capacité de recherche en sciences sociales, nous devons faire prendre conscience au grand public des avantages que présente la recherche en matière de politique sociale. Il nous faut démontrer que la recherche inductive fournit la base sur laquelle on peut s'appuyer pour élaborer des politiques gouvernementales, évaluer les programmes sociaux et surveiller l'efficacité des mesures visant des objectifs de société. Rien ne peut remplacer le dialogue entre chercheurs et décideurs sur les questions particulières qui les occupent. Comme ce dialogue ne se produira pas spontanément, il faut se doter des moyens voulus pour élargir la diffusion des résultats de recherches.

Un groupe de travail mixte du Conseil de recherches en sciences humaines et de Statistique Canada a été réuni par Ivan Fellegi, statisticien en chef, et Marc Renaud, président du CRSH. Son mandat est de présenter un ensemble de propositions ayant pour objet d'encourager la recherche quantitative, utilisant des données à grande échelle, sur des questions sociales et économiques d'envergure. La création de ce groupe vient s'ajouter à d'autres initiatives qui tendent à améliorer la recherche utilisant les statistiques sociales. Ainsi, le gouvernement fédéral a mis sur pied récemment le Projet de recherche sur les politiques (PRP), dont les responsables sont chargés de recommander un programme de recherche interministériel destiné à combler les lacunes que présentent les connaissances nécessaires à l'élaboration des politiques gouvernementales, et de superviser la mise en oeuvre de ce programme. L'équipe du PRP s'est associée avec le CRSH pour entreprendre des travaux sur les tendances de la société, dont l'objectif est de faire mieux comprendre les grands changements sociaux, de mettre en lumière leurs incidences sur la recherche axée sur les politiques, et de cerner les principaux domaines où il faudra effectuer des recherches dans les années à venir. En outre, le CRSH élabore en ce moment des programmes de recherche stratégique qui portent sur la cohésion sociale, sur l'économie et la société fondées sur les connaissances et sur la santé de la population. Ce sont là aussi des domaines prioritaires pour le PRP.

Le présent rapport recommande le financement des trois composantes d'un **Système d'analyse des statistiques sociales**, lesquelles visent principalement à augmenter le nombre de chercheurs qui font de la recherche quantitative en sciences sociales, à améliorer l'accès aux microdonnées sans compromettre la confidentialité, ainsi qu'à informer le public et à éveiller son intérêt pour les statistiques sociales. Le nouveau système viendrait compléter notre système statistique national, qui est déjà considérablement développé. Les composantes sont décrites séparément dans les trois prochaines parties du rapport. La dernière partie présente un plan d'exécution de même que le budget prévu pour la mise en oeuvre des propositions, et recommande un mode de coordination des activités interdépendantes qui suivraient l'adoption du rapport.

II. Accroître la capacité de recherche

La première composante du système que nous proposons vise à accroître le nombre de spécialistes qui font de la recherche quantitative en sciences sociales, et à hausser les compétences de ceux qui travaillent déjà dans le domaine. Elle réunit trois propositions distinctes : des **groupes de recherche et de formation**; un **programme de formation**, incluant un programme d'été, et enfin un **programme de bourses**, qui comprendrait des bourses de maîtrise, de doctorat et d'études postdoctorales ainsi que des bourses de recherche de niveau supérieur. Ces propositions sont décrites ci-dessous.

A. Les Groupes de recherche et de formation

Nous recommandons la mise sur pied d'un programme de groupes de recherche et de formation, qui seraient chargés de quatre fonctions interdépendantes : exécution de travaux de recherche quantitative sur des questions sociales; formation des chercheurs; large diffusion des résultats de la recherche, et rétroaction concernant l'amélioration des sources de données. Nous envisageons des groupes de taille moyenne, de préférence interuniversitaires et interdisciplinaires, qui s'attaqueraient à des problèmes de recherche d'envergure, centrés sur des données provenant de sources diverses. Le programme des Groupes de recherche et de formation serait doté d'un comité de sélection, incluant des chercheurs de l'étranger, qui appliquerait un processus de sélection stratégique, dynamique et en plusieurs étapes.

La réalisation d'un ensemble de recherches solides et diversifiées exige avant tout l'existence d'une communauté de spécialistes qui travaillent en collaboration, s'encouragent et se critiquent les uns les autres. Les communautés de chercheurs qui se concentrent sur l'étude de questions particulières ont une importance cruciale tout spécialement pour les étudiants diplômés et pour les chercheurs en début de carrière, parce qu'elles leur permettent d'acquérir de l'expérience et d'affiner leur jugement à un point qu'il serait difficile d'atteindre par des lectures ou par des cours universitaires. Le CRSH a pu observer le fonctionnement de différentes formes d'organisation, et il est arrivé à la conclusion que les milieux les plus productifs sont ceux où les chercheurs se retrouvent face à face pour travailler sur une question commune.

La pièce maîtresse de notre projet, c'est donc un programme d'appui aux équipes de recherche et de formation. Nous envisageons des groupes de taille moyenne, qui compteraient typiquement de huit à quinze chercheurs. Ce chiffre assure une masse critique, tout en évitant les problèmes de coordination que posent des groupes plus nombreux. Chacun des Groupes de recherche et de formation se consacrerait à l'étude d'un sujet assez vaste pour justifier un programme de coordination d'activités de recherche, de formation et de diffusion, et assez important pour attirer un nombre considérable de chercheurs. Idéalement, les Groupes seraient multidisciplinaires, et donneraient lieu à des partenariats entre des disciplines d'une même université, ou entre deux ou plusieurs universités. Ils feraient porter leurs efforts sur la recherche relative à un thème particulier. On trouvera à l'**annexe A** (page 21) une description de certains thèmes prometteurs sur lesquels on dispose dès à présent d'une quantité de données.

Il existe déjà des groupes de recherche sur les questions sociales au Canada, qui oeuvrent dans des domaines comme le vieillissement, l'éducation, l'immigration et le développement de l'enfant, et reçoivent une aide financière par l'intermédiaire des Centres nationaux d'excellence ou du programme des Grands Travaux de recherche concertée, dont s'occupe le CRSH. Les Groupes de recherche et de formation qui font l'objet de notre proposition, toutefois, se consacreraient à la recherche quantitative en sciences sociales, et accorderaient beaucoup plus d'importance au développement des compétences des nouveaux chercheurs ainsi qu'à l'enrichissement des compétences de toute l'équipe. Les bonnes équipes de recherche fournissent déjà d'excellentes occasions d'apprentissage non structuré grâce aux conversations, aux réunions et aux séminaires organisés, mais les Groupes de recherche et de formation seraient tenus de s'occuper plus systématiquement des besoins en matière de formation. Ils détermineraient ces besoins et définiraient des mesures concrètes visant à y répondre. Ils pourraient mettre sur pied des cours universitaires ou améliorer les cours existants, ou préparer des documents de travail sur les méthodes de recherche ou sur des aspects de la conception d'enquêtes. Les efforts

tendant à créer des stages qui jetteraient un pont entre les cours réguliers de niveau supérieur et la recherche appliquée auraient une importance primordiale.

On s'attendrait également à ce que les Groupes élaborent et mettent en oeuvre des plans de diffusion des résultats de leurs recherches, non seulement auprès d'auditoires spécialisés composés d'universitaires, mais aussi auprès des décideurs et du grand public. Ces plans supposeraient davantage qu'une fièvre de communication épisodique : ils appelleraient des efforts soutenus afin d'atteindre les groupes intéressés, et aboutiraient, dans plusieurs cas, à la création de partenariats avec des groupes de non-universitaires. Enfin, les Groupes de recherche et de formation feraient des observations aux responsables de la collecte de données, en particulier à Statistique Canada, à propos de la qualité des données que l'on réunit, de leur pertinence par rapport aux principaux problèmes de recherche, et des faiblesses qu'elles présentent. Nous pensons que ce serait un moyen d'accroître l'efficacité des futurs travaux de collecte de données, et d'encourager la participation continue des chercheurs à cette activité.

Avec le temps, les Groupes de recherche et de formation oeuvrant dans des domaines semblables ou connexes s'intégreraient à des réseaux virtuels – régionaux, nationaux ou internationaux. Ces réseaux permettraient aux chercheurs ne vivant pas à proximité d'établir des relations avec les Groupes, ou même d'en devenir membres. Les Groupes de recherche et de formation seraient incités à s'acquitter de leurs responsabilités en matière de formation, de diffusion et de rétroaction en mettant à profit d'autres éléments du Système canadien de recherche utilisant les statistiques sociales qui fait l'objet de notre proposition : le Programme d'été et les ateliers spécialisés qui seraient offerts par le Centre de formation; les Forums de recherche, qui donneraient l'occasion de rencontrer une diversité d'utilisateurs de la recherche, et les activités de communication qui seraient élaborées au CRSH.

Nous croyons que la plupart des Groupes de recherche et de formation se consacreraient à l'analyse d'enquêtes à grande échelle de Statistique Canada. Le CRSH et Statistique Canada encourageraient la recherche utilisant des données complexes, longitudinales et multiniveaux, et la recherche comparée, surtout celle qui a recours à des bases de données multinationales. Les Groupes de recherche et de formation pourraient aussi planifier et mener leurs propres enquêtes, ou bien de façon autonome, comme ce serait le cas pour des études sur les élections au Canada, ou bien à titre de participants à une enquête sociale d'envergure internationale. La réalisation de telles enquêtes à l'extérieur de Statistique Canada, ou en partenariat avec cet organisme, donnerait plus d'occasions de former des étudiants à la planification, à la conception et à l'exécution des enquêtes sociales. Cette formule serait avantageuse tant pour Statistique Canada que pour l'industrie de la recherche-sondage.

Les demandes relevant du programme des Groupes de recherche et de formation feraient l'objet d'un examen par les pairs, et les subventions seraient accordées pour une période initiale de cinq ans, avec possibilité de renouvellement. Le budget typique prévoirait les fonds nécessaires pour la rémunération des étudiants engagés comme assistants de recherche et pour celle des titulaires de bourses d'études postdoctorales, pour l'achat d'ordinateurs et de logiciels, pour les déplacements et, dans certains cas, pour la collecte et l'archivage des données. Une aide financière serait également accordée pour la rémunération de membres de l'équipe qui seraient recrutés ailleurs que chez les étudiants; les Groupes auraient probablement besoin, en particulier, de professionnels capables de préparer et de documenter les fichiers de données complexes qui sont indispensables à l'exécution d'analyses quantitatives élaborées, de fournir un appui continu aux étudiants des cycles supérieurs qui entreprennent l'analyse de données, et de résoudre les problèmes de fonctionnement des ordinateurs et des logiciels.

Nous recommandons que le programme des Groupes de recherche et de formation soit doté d'un comité de sélection unique et distinct, qui compterait parmi ses membres des spécialistes de l'étranger possédant une connaissance approfondie de la recherche quantitative en sciences sociales, y compris de la recherche axée sur les politiques. Le comité serait chargé d'élaborer un processus de sélection à la fois détaillé et formateur, c'est-à-dire qui inciterait à formuler de

nouvelles propositions et à établir des liens avec d'autres groupes ainsi qu'avec le milieu de l'élaboration des politiques. Nous estimons qu'il faut un comité de sélection distinct parce qu'un tel comité serait en mesure de comprendre les exigences particulières de la recherche quantitative en sciences sociales, et d'équilibrer les critères, les préférences et les objectifs qui s'appliqueraient aux Groupes de recherche et de formation. Ce comité aurait également pour fonction d'inscrire l'activité des Groupes dans le cadre du Système canadien de recherche utilisant les statistiques sociales, et de veiller à ce que les Groupes fassent des progrès continus, surtout lorsqu'ils se trouvent encore aux premières étapes de leur développement.

B. Améliorer les compétences : le Programme de formation et le Programme d'été

> Nous recommandons la mise sur pied d'un programme de formation, incluant un programme national d'été, qui traiterait de divers sujets se rapportant aux méthodes quantitatives et à l'utilisation des microdonnées recueillies au Canada. Ce programme s'adresserait aux analystes de politiques, aux chercheurs universitaires ainsi qu'aux étudiants des trois cycles. Il fournirait aussi un appui aux bibliothécaires responsables des données.

Il convient d'établir, en complément des Groupes de recherche et de formation, un programme de cours de brève durée sur l'analyse des données sociales, semblable à ce que les Américains et les Britanniques ont mis sur pied à Ann Arbor et dans l'Essex, mais plus décentralisé et mieux adapté aux données recueillies au Canada ainsi qu'aux préoccupations de nos chercheurs. Notre but n'est pas d'offrir aux étudiants des trois cycles un substitut des cours universitaires, mais de présenter une série de cours de niveau intermédiaire sur l'analyse des données, et de cours avancés sur l'analyse des données longitudinales et des données multiniveaux. Ces cours feraient une large place à l'application des techniques statistiques, et utiliseraient des microdonnées réunies au Canada.

La création d'un tel programme est nécessaire pour plusieurs raisons. Tout d'abord, les étudiants de la fin du baccalauréat et les étudiants diplômés qui suivent une formation en statistique sont souvent incapables de mettre à profit l'enseignement théorique qu'ils ont reçu pour appliquer les techniques à des données réelles. Dans bien des cas, les exemples qui figurent dans les manuels se rapportent à des questions sociales concernant les États-Unis, et utilisent des ensembles de données de petite taille et artificiels. Bon nombre d'étudiants ont besoin par ailleurs d'une aide individualisée pour apprendre les rudiments de la gestion des données : comment extraire des données d'un CD-ROM, créer des fichiers de synthèse, choisir des cas, trier les données, etc. En général, ils ont également besoin de connaissances sur des sujets assez simples qui sont rarement abordés dans les cours, comme la construction d'échelles, la création de variables composites, le traitement des données manquantes et l'utilisation des coefficients de pondération des plans de sondage. Quelques cours de base utilisant un ou deux ensembles de données recueillies au Canada pourraient susciter un grand enthousiasme pour une carrière en statistiques sociales. En outre, lorsque des étudiants suivent ce genre de cours avec d'autres chercheurs universitaires et des analystes de politiques, ils voient eux-mêmes à quels types de problèmes peut s'appliquer l'analyse des microdonnées. Les participants apprennent beaucoup de façon fortuite et il se forme d'importantes relations de travail, souvent interdisciplinaires.

Les chercheurs de niveau avancé ont besoin quant à eux d'être formés à l'utilisation des méthodes statistiques complexes qui conviennent à l'analyse des ensembles de données longitudinales et multiniveaux. Une telle formation est essentielle si nous voulons exploiter toute

la richesse des enquêtes qui ont été mises au point récemment. La clientèle des cours serait constituée principalement de chercheurs universitaires, d'une part, et d'analystes de politiques travaillant dans la fonction publique ou dans divers autres organismes, d'autre part. Ces cours devraient être donnés dans un lieu central, parce qu'il ne se trouve habituellement pas, dans une même université, une masse critique de chercheurs souhaitant recevoir une formation aussi avancée.

Enfin, comme les bibliothécaires responsables des données jouent un rôle essentiel en facilitant l'accès aux bases de données et leur utilisation, il convient de prendre les mesures voulues pour favoriser leur perfectionnement. On contribuerait du même coup à la formation des étudiants non diplômés, parce que les bibliothécaires responsables des données peuvent aider les professeurs à se servir dans leurs cours de données recueillies au Canada, et qu'ils peuvent en outre aider les étudiants à faire démarrer des projets qui comportent l'analyse de données.

Il est logique que les efforts visant à améliorer la formation à l'analyse de données quantitatives au Canada soient centrés sur un programme d'été, inspiré du modèle que proposent les programmes d'été permanents qui remportent un grand succès dans le Michigan et l'Essex. Nous envisageons une série de cours dans une ou plusieurs villes, normalement sur une période de trois à quatre semaines. On trouvera à l'**annexe B** (page 27) des idées quant aux cours qui pourraient être offerts. Notre programme différerait cependant des programmes existant aux États-Unis et en Grande-Bretagne sous trois aspects importants : premièrement, il porterait sur des questions sociales qui se posent au Canada et ferait usage de données recueillies au pays; deuxièmement, il créerait un environnement social et intellectuel pour les étudiants diplômés, les membres du corps professoral et les chercheurs du secteur public qui s'intéressent à la recherche quantitative sur des questions concernant le Canada; troisièmement, il fournirait les moyens de lancer des activités de suivi, par exemple la mise sur pied d'un réseau électronique ou l'échange de documents de recherche. Il donnerait aussi aux chercheurs l'occasion d'entrer en relations avec des employés de Statistique Canada qui travaillent directement à la gestion de certaines bases de données.

Rien ne nous oblige à restreindre ces activités à la saison estivale. Le Centre de formation devrait favoriser l'organisation d'activités connexes, en matière de formation et de consultation, pendant toute l'année. Celles-ci pourraient inclure des ateliers visant la planification de nouvelles enquêtes, des séminaires ayant pour objet d'appuyer la recherche sur un sujet déterminé ou à partir d'un certain ensemble de données, et des ateliers sur les techniques statistiques et les méthodes d'enquête. Les installations pourraient être mises à la disposition des ministères, des fondations, des entreprises, des syndicats et des ONG pour des séances de consultation et des séminaires sur des questions auxquelles les chercheurs doivent s'attaquer d'urgence. Des activités de cette nature se tiennent déjà en réponse à des besoins particuliers, mais un centre ayant pour mandat de les faciliter réduirait de façon considérable les efforts que doivent actuellement fournir les chercheurs afin d'offrir de tels programmes de formation.

Nous proposons de commencer par un programme organisé par Statistique Canada, et tirant parti des installations d'enseignement dont cet organisme dispose à Ottawa. Statistique Canada non seulement présente un bon milieu d'apprentissage, déjà doté des outils nécessaires, mais se trouve aussi en mesure d'offrir des services en français et en anglais, ce qui constitue un immense avantage. Dès le début, Statistique Canada travaillerait en étroite collaboration avec les universitaires par l'intermédiaire d'un comité consultatif. Une fois le programme bien implanté, certains cours pourraient se donner dans diverses universités. Par la suite, le programme lui-même pourrait relever d'une ou plusieurs universités, Statistique Canada conservant un rôle de partenaire. Il faudrait en outre s'efforcer de conclure avec des universités une entente visant l'agrément des cours, ce qui serait sans doute utile pour beaucoup d'étudiants.

Nous recommandons qu'une aide financière soit accordée pour lancer le programme. Cette aide devrait comprendre le traitement du directeur, les frais de bureau, les frais de déplacement, le traitement des enseignants, ainsi que les frais de voyage et les honoraires des membres du

comité consultatif. Le Centre serait financé principalement par les droits de scolarité et le parrainage de certaines activités, comme les ateliers sur un ensemble de données en particulier. Le programme devrait chercher à s'autofinancer à long terme. Par ailleurs, afin d'attirer des étudiants des cycles supérieurs à un programme d'été, il faudrait un appui du CRSH pour les frais de déplacement, le logement et les droits de scolarité. Un appui semblable pourrait aussi être offert aux jeunes professeurs. Il conviendrait de s'adresser à ce propos au CRM et au CRSNG, qui pourraient soutenir des étudiants diplômés dans des domaines qui bénéficieraient du programme.

C. Les programmes de bourses

> Nous recommandons que le CRSH mette sur pied un programme ciblé visant à fournir une aide financière aux étudiants et aux jeunes chercheurs en statistiques sociales du niveau de la maîtrise, du doctorat ou des études postdoctorales, et un programme de bourses de recherche de niveau supérieur en statistiques sociales.

Afin de consolider le milieu de la recherche à long terme, nous devons faire en sorte que de jeunes chercheurs soient attirés par une carrière en statistiques sociales. Cette discipline doit rehausser son image de marque, et se faire bien connaître des étudiants diplômés ou de niveau postdoctoral qui ont à décider de l'orientation de leurs recherches. Il conviendrait, à cette fin, d'offrir aux intéressés une aide financière qui leur soit spécialement destinée.

Bourses de maîtrise, de doctorat et d'études postdoctorales. Ce programme de bourses serait distinct des programmes de bourses réguliers du CRSH, mais les conditions d'admissibilité, les règlements et le montant de l'aide accordée seraient les mêmes. Le concours serait réservé aux étudiants qui font de la recherche quantitative en sciences sociales, mais le moyen de prouver le respect de ce critère varierait selon le niveau d'études. Les candidats aux bourses de maîtrise seraient tenus de suivre des cours de deuxième cycle leur apportant des compétences en mathématiques et en analyse, et de rédiger soit un article de synthèse ou une thèse sur un sujet se rapportant à la recherche quantitative en sciences sociales. Les candidats aux bourses de doctorat ou d'études postdoctorales devraient soumettre un projet de recherche quantitative.

Comme il s'agirait d'un programme ciblé, l'évaluation des demandes exigerait la création d'un comité de sélection distinct, composé de spécialistes. Il faudrait aussi s'engager explicitement à réaliser l'équilibre des sexes. Nous recommandons en conséquence que toutes les demandes relatives à ces bourses soient soumises directement au CRSH.

Bourses de recherche de niveau supérieur en statistiques sociales. Le comité recommande que *deux* bourses de recherche de niveau supérieur soient accordées tous les deux ans à des universitaires qui contribuent de façon remarquable à la recherche et à la formation dans le domaine des statistiques sociales concernant le Canada. Les bourses, d'une valeur de 100 000 $ par année, seraient attribuées pour une période de cinq ans; on exigerait des bénéficiaires qu'ils consacrent au moins 80 p. 100 de leur temps à la recherche et à la formation. Le renouvellement serait consenti sous réserve que les bénéficiaires continuent d'exécuter des travaux d'une haute valeur scientifique. Vu leur caractère prestigieux, ces bourses auraient sans doute pour effet d'accélérer l'avancement de la carrière des récipiendaires. Elles seraient conçues suivant le modèle de celles qui sont accordées par l'Institut canadien des recherches avancées. Les bourses de cet organisme ont décidé certains de nos meilleurs chercheurs à rester au Canada, et leur ont permis d'élaborer des programmes de recherche d'envergure.

Nous recommandons que le comité de sélection soit le même que pour les autres bourses proposées ci-dessus. Les candidatures pourraient être soumises au CRSH soit directement, soit par l'intermédiaire des universités. La demande devrait expliquer brièvement de quelle façon l'exécution du projet de recherche favoriserait l'avancement des statistiques sociales et la formation de jeunes chercheurs. Le comité de sélection fonderait ses décisions sur l'étude des dossiers d'activité de recherche, de même que sur l'examen des programmes proposés en matière de recherche et de formation.

Stages. Nous estimons par ailleurs que les jeunes chercheurs doivent être plus nombreux à acquérir une expérience de la recherche axée sur les politiques, en vue notamment de mettre leurs nouvelles compétences au service des sciences sociales faisant appel aux données quantitatives. Cependant, bien que nous soyons conscients des immenses avantages que l'on pourrait tirer d'un programme de stages, nous n'avons pu formuler de suggestions précises à ce sujet, en l'absence d'un programme général au CRSH. Nous devons donc nous borner à recommander que le nouveau programme de stages proposé dans le scénario d'innovation du Conseil tienne compte du besoin particulier d'expertise en statistiques sociales qui est signalé dans le présent rapport, et que l'on prévoie les ressources appropriées pour répondre à ce besoin. Nous croyons en outre que Statistique Canada, en vue de favoriser l'établissement du Système d'analyse des statistiques sociales, devrait s'employer davantage à promouvoir ses propres programmes de stages, et inciter les ministères à faire de même.

III. Faciliter l'accès aux données

Pour que les Canadiens bénéficient pleinement des ressources considérables dont nous disposons en matière de données, les chercheurs et les analystes doivent pouvoir consulter ces dernières sans trop de difficultés. La recherche est indispensable pour convertir les données en information utilisable, et ne saurait être menée à bonne fin sans un accès adéquat. L'accès aux données présente bien des dimensions. Cela peut supposer que l'on veille à ce que les utilisateurs soient suffisamment au courant des principales statistiques que renferment les sources de données, et à ce qu'ils puissent les consulter. Cela peut supposer aussi que l'on permette la consultation des fichiers de microdonnées à usage public dont l'étude est nécessaire pour une analyse plus complexe, ou même, à un autre niveau, que l'on permette la consultation de microdonnées détaillées qui n'ont pas été soumises à une présélection (données confidentielles). La question de l'archivage est également pertinente dans ce contexte, puisqu'il deviendra impossible de consulter les sources de données, dans un avenir plus ou moins lointain, si on ne prend pas soin de les archiver convenablement dès leur création.

L'Initiative de démocratisation des données a grandement amélioré l'accès aux sources de données à usage public qui servent aux enquêtes sur les ménages. Le Groupe de travail mixte estime qu'elle a eu des effets extrêmement positifs, et recommande ci-dessous des façons de l'élargir. D'autres recommandations concernent un appui pour l'archivage des données qui sont créées par des groupes de recherche ou dans le cadre d'initiatives ne relevant pas de Statistique Canada.

La principale question qui nous préoccupe ici, cependant, se rapporte aux modalités suivant lesquelles les chercheurs pourraient avoir accès aux données qui sont tenues pour confidentielles en conformité avec la *Loi sur la statistique*. Il s'agit essentiellement de déterminer de quelle façon on pourrait permettre la consultation de ces données tout en assurant la confidentialité des réponses, comme le prescrit la *Loi*. En effet, même si la confidentialité est la pierre angulaire du système statistique, il faut reconnaître que, pour exploiter à fond les nouvelles sources de données qui sont créées aujourd'hui, les chercheurs ont besoin d'avoir une certaine forme d'accès aux données détaillées. Nous examinons ici divers moyens d'atteindre ces

objectifs, mais notre proposition la plus importante a trait à la création de centres d'accès aux données de recherche.

A. Les Centres d'accès aux données de recherche

> Nous recommandons que le CRSH et Statistique Canada établissent conjointement un système national de centres d'accès aux données de recherche, qui permettraient aux chercheurs de consulter les microdonnées détaillées pour les besoins de leurs travaux, tout en assurant le respect des dispositions de la Loi sur la statistique en matière de protection de la confidentialité.

Les chercheurs ont de plus en plus besoin de consulter les microdonnées détaillées afin d'exécuter des travaux dans de nombreux domaines, notamment la santé, le développement de l'enfant, le revenu et la sécurité économique, l'adaptation de la main-d'oeuvre et l'évolution du milieu de travail. Paradoxalement, à mesure qu'augmentent la qualité et la complexité de l'information recueillie, l'accès aux microdonnées risque de se trouver réduit en raison des exigences de confidentialité. De fait, il est crucial de protéger la confidentialité : non seulement on s'engage envers les citoyens à ne pas divulguer les renseignements qu'ils fournissent, mais la sécurité des données est d'une importance vitale si l'on veut continuer d'inspirer, tant aux milieux politiques qu'à la population en général, la confiance voulue pour mener à bien un large éventail d'activités de collecte de données.

Depuis le début des années 70, Statistique Canada a mis les microdonnées (les dossiers personnels) à la disposition des chercheurs, tout en protégeant leur caractère confidentiel, grâce à la production de fichiers de microdonnées « à usage public » pour les enquêtes effectuées sur les ménages. Afin d'éviter la communication de dossiers contenant une information qui pourrait permettre d'identifier les répondants, il faut supprimer certaines données – soit que l'on omette certains éléments des dossiers personnels ou que l'on « comprime » des catégories de réponses, de façon qu'il ne soit pas possible d'identifier des répondants qui présentent des caractéristiques inhabituelles. Cette solution n'est cependant plus adéquate, pour deux raisons. Premièrement, beaucoup de nouvelles enquêtes longitudinales ont une structure de données qui empêche la création même de fichiers de microdonnées à usage public : il devient en effet de plus en plus difficile de constituer un fichier où la confidentialité soit assurée, dans les cas où les données longitudinales sont cumulées pendant un certain nombre d'années. Deuxièmement, les chercheurs qui se servent de théories complexes et de techniques statistiques poussées se rendent compte de plus en plus que la suppression d'information, dans les fichiers à usage public qu'on met traditionnellement à leur disposition, limite sérieusement leur analyse. Si on ne règle pas cette question de l'accès aux données, non seulement un bon nombre de sources de données nouvelles et novatrices se verront sous-utilisées, et des travaux de grande valeur seront perdus, mais les chercheurs se tourneront peut-être de plus en plus vers des données provenant d'autres pays, en particulier des États-Unis, où l'on a trouvé une meilleure solution au problème d'accès. Or, il arrive souvent que les conclusions auxquelles permettent d'arriver les données recueillies chez nos voisins du Sud ne s'appliquent pas à la situation canadienne.

Un des moyens de résoudre la question serait de faire prêter serment aux chercheurs, en vertu de la *Loi sur la statistique*, comme « réputés être des personnes employées » par Statistique Canada, ce qui leur permettrait d'effectuer des recherches dans les fichiers de microdonnées détaillées. La *Loi sur la statistique* prévoit un tel statut de personne réputée employée, à des conditions dont nous parlerons plus bas. Même une fois que l'on a prêté serment, toutefois, demeure la question de l'accès matériel aux données. Nous proposons à ce sujet que soient

créées au Canada un certain nombre d'installations où la consultation serait possible. Ces installations porteraient le nom de **Centres d'accès aux données de recherche**, et s'inspireraient d'un programme semblable qui a été mis sur pied par le U.S. Census Bureau. Voici un résumé de notre proposition, qui est exposée en détail à l'**annexe C** (page 29).

Nous prévoyons, à plus ou moins longue échéance, la création de plusieurs centres d'accès aux données de recherche, répartis dans de grandes villes et dans des campus universitaires de tout le pays. Les Centres fourniraient des installations sûres pour la conservation des données confidentielles. D'un point de vue juridique ce seraient des bureaux de Statistique Canada, et un employé de cet organisme travaillerait sur place. Ces centres assureraient le fonctionnement d'un programme de recherche administré par le CRSH et par Statistique Canada. Les chercheurs dont le projet aurait été accepté dans le cadre du programme, après avoir prêté serment en vertu de la *Loi sur la statistique*, auraient accès aux fichiers de microdonnées confidentielles qui seraient conservés dans les Centres. Un comité relevant conjointement du CRSH et de Statistique Canada s'occuperait de la sélection des demandes de participation, en fondant ses décisions sur la valeur du projet de recherche. Ce comité de sélection se composerait en grande partie de chercheurs universitaires, auxquels se joindraient des représentants d'autres communautés de chercheurs et de Statistique Canada. La valeur du projet serait le tout premier critère de sélection (avec la question des ressources). Les demandes pourraient venir de chercheurs appartenant à des universités, à des instituts de recherche, à des organismes gouvernementaux ou à d'autres organismes de recherche.

La *Loi sur la statistique* régit l'accès aux données confidentielles, et ses dispositions doivent faire partie intégrante de tout programme des Centres d'accès aux données de recherche. Les chercheurs peuvent devenir des personnes réputées employées par Statistique Canada, et avoir accès aux données confidentielles, à condition de rendre « des services spéciaux au ministre » (c.-à-d. à Statistique Canada). Comme Statistique Canada, de toute évidence, ne dispose pas des ressources humaines indispensables afin d'effectuer toutes les recherches nécessaires à l'exploitation des sources de données, cet organisme est justifié, en vue d'atteindre son objectif, d'engager d'autres chercheurs qui fourniront ce service. Dans ces circonstances, les travaux à effectuer seraient semblables à ceux que Statistique Canada lui-même exécuterait normalement. Cette condition peut sembler restrictive, mais comme Statistique Canada réalise en fait un large éventail de travaux de recherche quantitative, elle ne devrait pas susciter de difficultés.

Le programme de recherche relevant des Centres d'accès aux données aurait d'abord pour résultat la production de rapports de recherche par les utilisateurs. Ces rapports seraient soumis au processus habituel d'examen par les pairs, qui serait géré par le comité de sélection. On s'attend à ce que la grande majorité d'entre eux soient inclus par la suite dans la série de documents de recherche associée au programme. Comme celui-ci serait dirigé conjointement par le CRSH et par Statistique Canada, les documents ne pourraient pas renfermer de commentaires directs sur les politiques ou les programmes, étant donné que de telles observations ne font pas partie du mandat de Statistique Canada (voir l'annexe C, à la page 29, pour des détails). Une fois passée cette première étape, les chercheurs auraient toute liberté pour faire paraître un rapport révisé (ou le rapport original) dans une revue spécialisée ou dans toute autre publication. On ne se préoccuperait pas, dans ce cas, de savoir si le texte contient des observations relatives aux politiques : les auteurs seraient libres d'y incorporer tous les éléments qu'ils jugeraient nécessaires. Nous proposons que les droits d'auteur relatifs au rapport de recherche original appartiennent aux chercheurs. Statistique Canada se réserverait le droit de publier le texte original, et de s'assurer que toutes les publications respectent les règles de confidentialité.

Pourvu que les installations répondent à des normes de sécurité comparables à celles qui sont en vigueur à Statistique Canada, les Centres d'accès aux données de recherche pourraient être situés dans des universités ou dans des établissements de recherche non gouvernementaux. Pour des raisons de commodité, toutefois, certains d'entre eux pourraient être logés dans des bureaux régionaux de Statistique Canada. L'avantage de choisir des locaux à l'extérieur de Statistique Canada, c'est que l'établissement hôte aurait la chance de développer très fortement

sa capacité de recherche inductive. Les chercheurs qui s'orientent dans cette direction seraient attirés par un établissement pourvu d'un centre d'accès aux données, ce qui permettrait à l'université ou à l'institut d'élaborer un programme solide. Il pourrait être approprié, toutefois, de lancer le programme par un projet pilote, dans un bureau régional de Statistique Canada. Nous proposons par ailleurs que les Centres d'accès aux données de recherche commencent par autoriser la consultation d'enquêtes et de fichiers de données administratives excluant les documents relatifs aux entreprises; l'accès à ces derniers soulève en effet des questions spécifiques, qu'il faudrait régler avant d'intégrer au programme les documents en cause.

On peut envisager au moins deux méthodes de financement pour les Centres d'accès aux données de recherche. Selon la première, conçue d'après le modèle de l'Initiative de démocratisation des données, les Centres feraient l'objet d'un « financement global » couvrant toutes leurs activités, dont le coût serait partagé, suivant une formule à déterminer, entre le CRSH, Statistique Canada, l'université ou l'institut hôte et d'autres d'organismes, comme des ministères ou des instituts de recherche. Si, comme nous le pensons, les centres étaient utilisés principalement par des chercheurs universitaires, la plus grande partie de ce financement global viendrait du CRSH. Les chercheurs n'auraient pas à payer de frais directs, et l'accès aux données ne dépendrait pas de l'attribution d'une aide financière, mais plutôt de la décision du comité de sélection. Une autre possibilité serait de prévoir un financement global pour une partie du coût de fonctionnement des Centres, et d'imposer aux chercheurs des droits d'utilisation. Les chercheurs universitaires dont le projet serait approuvé par le comité recevraient du CRSH une subvention couvrant ces droits; les chercheurs d'autres organismes auraient à payer des droits semblables. Quoi qu'il en soit, il faudrait prendre soin de faciliter l'accès aux données pour les étudiants des cycles supérieurs et les chercheurs de niveau postdoctoral.

Créer un service de consultation des données à distance. Outre la possibilité d'autoriser des personnes réputées employées à consulter les données confidentielles dans les Centres d'accès aux données de recherche, on peut envisager au moins deux autres solutions au problème d'accès aux données. L'une d'elles consisterait à doter Statistique Canada d'un service de consultation à distance. Les chercheurs soumettraient à l'organisme leurs travaux informatiques, et le personnel de Statistique Canada ferait l'extraction des données voulues à partir des fichiers de microdonnées confidentielles. Les résultats seraient contrôlés par des employés de Statistique Canada, qui vérifieraient qu'on ne divulgue pas de données confidentielles. Il s'agirait là d'un complément aux Centres d'accès aux données de recherche, et non d'une solution de remplacement. Le service de consultation à distance serait utile aux chercheurs exécutant des projets de petite envergure, ou à ceux qui n'habiteraient pas à proximité des Centres d'accès aux données de recherche. Bien entendu, il n'offrirait cependant pas la souplesse ni les possibilités d'interaction rapide qu'exigent beaucoup de travaux de recherche, et que les Centres d'accès seraient en mesure d'assurer. Un comité de Statistique Canada a été constitué pour formuler une proposition à ce sujet; on trouvera à l'**annexe D** (page 35) les grandes lignes de ses conclusions.

Obtenir des répondants l'autorisation de communiquer les données aux utilisateurs des Centres d'accès. La *Loi sur la statistique* inclut une disposition relative au partage de données avec des utilisateurs. Si l'on obtient l'autorisation voulue auprès des répondants au moment de la collecte des données, les microdonnées confidentielles peuvent être communiquées à certains utilisateurs. Il serait donc possible de partager des données avec un institut créé par le CRSH, ou avec un autre organisme. Statistique Canada fournirait les microdonnées brutes (sur les répondants qui auraient accordé leur autorisation, comme le font la plupart) à l'institut, qui s'engagerait évidemment, en vertu de dispositions contractuelles, à en préserver la confidentialité.

Ce mécanisme est assez simple. Le processus de sélection et de contrôle relèverait entièrement de l'institut, Statistique Canada ne faisant guère plus que fixer les conditions visant à protéger la confidentialité. Le défaut de cette méthode, c'est qu'elle peut s'appliquer uniquement aux données qui seront réunies dans les années à venir, et à certains ensembles de données. Les ensembles qui ont déjà été constitués ne pourraient pas faire l'objet d'une telle entente. Pour

cette raison, le Groupe de travail mixte estime que cette idée mérite d'être approfondie pour l'avenir, mais qu'elle ne résoudra pas les questions actuelles. À long terme, ce pourrait être une excellente formule pour le partage de certains ensembles de données, et on devrait l'envisager sérieusement une fois que les mesures proposées dans le présent rapport auront commencé à porter leurs fruits.

B. L'archivage et la documentation des données

> Nous recommandons que le CRSH permette aux chercheurs qui ont créé des fichiers de données avant que le coût de la documentation et de l'archivage soit couvert par les subventions de cet organisme de présenter une demande d'aide financière visant l'exécution de tels travaux, en ce qui a trait aux fichiers de données les plus importants. Nous recommandons aussi que le CRSH ou les responsables de l'Initiative de démocratisation des données envisagent la mise sur pied d'un système d'archivage national, pour les ensembles de données quantitatives concernant les sciences sociales qui ont été créés à l'extérieur de Statistique Canada.

Il arriva fréquemment que d'autres organismes que Statistique Canada produisent de vastes bases de données de recherche quantitative. Le CRSH finance des initiatives qui peuvent générer ce type de fichiers. Les plus importantes de ces sources de données devraient être archivées en vue d'une utilisation à venir, et les bases de données devraient être établies et entretenues d'une façon qui les rende facilement accessibles et transportables. Si de telles mesures ne sont pas prises, les chercheurs qui n'ont pas travaillé avec ces données dès le départ sont incapables d'y avoir accès, en particulier lorsqu'il y a quelque temps que le fichier a été créé. À l'heure actuelle, on n'accorde pas une aide financière suffisante pour l'archivage d'importantes bases de données de recherche qui ont été constituées il y a un certain nombre d'années. Certes, les dépenses qu'entraînent la documentation des données et la création de fichiers à des fins de diffusion sont aujourd'hui autorisées pour les travaux de collecte financés par le CRSH. La difficulté demeure, toutefois, pour les bases de données qui sont établies sans l'appui du CRSH, ou avec un appui partiel sur une assez longue période, et pour celles qui ont été mises sur pied avant qu'une aide soit allouée pour l'archivage. C'est le cas notamment de l'ensemble de données sur le vieillissement au Manitoba, de l'Étude sur la santé et le vieillissement au Canada et de l'Enquête canadienne sur la fécondité. Nous croyons comprendre qu'il n'est pas possible, en ce moment, de demander une subvention pour présenter les données sous une forme qui en permette la diffusion. Nous recommandons en conséquence que le CRSH élargisse un de ses programmes ou mette sur pied un nouveau programme qui accorde ce genre de subvention. Les demandeurs devraient faire ressortir que les données n'existent pas actuellement sous une forme qui en favorise la communication ou l'archivage, indiquer de quelle façon ils comptent s'y prendre pour préparer les données, décrire le produit qu'ils entendent obtenir et montrer que les données présentent une importance et un intérêt tels qu'elles seront véritablement utilisées si elles sont archivées et diffusées.

C. L'Initiative de démocratisation des données

> Nous recommandons que les responsables de l'Initiative de démocratisation des données soient invités à soumettre des propositions visant à poursuivre et à étendre la mission de ce programme, qui est de rendre les données et l'information sur la recherche facilement accessibles pour les besoins de la recherche universitaire et de la formation des étudiants non diplômés; ces propositions pourraient concerner la formation des bibliothécaires responsables des données et l'élargissement des réseaux qui favorisent la collaboration entre eux et avec les professeurs enseignant au premier cycle, ainsi que la préparation, la diffusion et l'échange de matériel didactique.

L'Initiative de démocratisation des données (IDD) a été lancée il y a quelques années par la Fédération canadienne des sciences humaines et sociales et Statistique Canada, de concert avec un groupe d'établissements postsecondaires et de ministères fédéraux. Elle a remporté un immense succès, et a grandement facilité l'accès aux fichiers de données à usage public pour les chercheurs des universités et des collèges. Beaucoup plus de chercheurs commencent à s'intéresser aux données quantitatives. Une nouvelle génération de bibliothécaires responsables des données ont été nommés et formés dans la plupart des établissements; ils ont appris à travailler en étroite collaboration les uns avec les autres afin d'accroître les possibilités de consultation aussi bien des données que de l'information nécessaire pour les analyser. Avant l'IDD, très peu d'établissements étaient pourvus d'un centre de données; leur nombre a considérablement augmenté depuis. On a en outre élaboré des logiciels qui fournissent des métadonnées sur les sources de données existantes, et qui aident à extraire les renseignements : c'est dire que les étudiants et les chercheurs relativement inexpérimentés peuvent maintenant obtenir, facilement et rapidement, les données quantitatives dont ils souhaitent se servir pour leurs travaux d'analyse; on alimente ainsi leur intérêt, ce qui prépare la voie à d'autres études utilisant les statistiques sociales.

Depuis le lancement de l'IDD, un bon nombre de bibliothécaires responsables des données ont aussi travaillé en collaboration avec les professeurs de sciences sociales qui enseignent aux étudiants de premier cycle la logique et la pratique de l'analyse des données, ainsi que les méthodes s'y rapportant. Beaucoup de ces professeurs ont notablement élargi la place que tiennent dans leur enseignement les travaux pratiques utilisant les données. Ils peuvent d'ailleurs envoyer aux bibliothécaires leurs étudiants, de premier cycle ou des cycles supérieurs, pour obtenir de l'aide afin de consulter les données et la documentation qui les accompagne.

L'Initiative de démocratisation des données fait actuellement l'objet d'un réexamen, à l'approche du terme de sa phase initiale de cinq ans. L'Initiative est menée par un organisme autonome, pourvu d'un personnel de gestion et de son propre conseil d'administration, et il n'appartient pas au Groupe de travail mixte de défendre sa cause à la place de ses dirigeants. Cela dit, nous devons faire observer à quel point il est important que cette entreprise reçoive, de Statistique Canada comme du CRSH, un appui aussi bien moral que financier. Il faut bien entendu préserver la raison d'être de l'IDD, qui est de diffuser les données à usage public auprès des chercheurs universitaires; mais l'organisme a assumé d'autres rôles qu'il convient de reconnaître, d'encourager et d'élargir. Premièrement, la formation des bibliothécaires responsables des données, qui sont devenus des intermédiaires indispensables entre les sources de données et les utilisateurs, devrait être enrichie, et les normes d'exercice de la profession devraient être

sanctionnées par les employeurs comme par les employés. En outre, les réseaux qui favorisent la collaboration entre les bibliothécaires devraient être étendus et renforcés.

Deuxièmement, l'IDD a toutes les caractéristiques voulues pour jouer un rôle à la fois plus grand et plus systématique, en collaboration avec les établissements d'enseignement postsecondaire ainsi qu'avec la Fédération canadienne des sciences humaines et sociales, Statistique Canada et le CRSH, afin d'améliorer l'enseignement que l'on donne au premier cycle en ce qui a trait aux méthodes employées en statistiques sociales. Elle peut favoriser en effet les échanges entre les professeurs du premier cycle, de même qu'avec les bibliothécaires responsables des données, à propos de techniques d'enseignement améliorées, soit par écrit – sur support papier ou électronique –, soit au moyen d'ateliers, par exemple au congrès annuel des sciences sociales. Elle peut aussi améliorer le matériel didactique qui est mis à la disposition des professeurs, que ce soit sous forme de données brutes ou à demi traitées, de documentation sur les données et les concepts, d'articles et de rapports de recherche, ou de suggestions de travaux pratiques.

Troisièmement, le travail accompli par les responsables de l'IDD, aussi bien pour l'ensemble du pays que dans les centres de données locaux et auprès des réseaux, peut être vu comme préparatoire à la création d'une archive des données canadiennes. De telles archives existent déjà dans la plupart des pays avancés, et se sont révélées des outils de recherche incomparables. La mise sur pied d'archives de ce type, au niveau national, n'est pas prévue à l'heure actuelle, mais cela peut changer, et le travail effectué par l'IDD afin de produire des métadonnées et de la documentation de meilleure qualité a de fortes chances de nous orienter dans cette direction. Il convient donc de l'appuyer fortement. D'ailleurs, la proximité des bibliothécaires responsables des données par rapport aux utilisateurs va contribuer à rendre le matériel plus convivial pour ces derniers. Nous pouvons prévoir que les centres de données situés dans les universités réussiront même à préserver et à diffuser la précieuse information que l'on réunit sur les procédures concrètes de recherche lorsqu'on utilise les données quantitatives. En d'autres termes, une archive de données pourrait très bien se constituer à la faveur des activités qui auront cours durant les quelques premières années qui suivront le lancement du Système canadien de recherche utilisant les statistiques sociales; le CRSH et Statistique Canada devraient surveiller l'évolution des choses, et se tenir prêts à soutenir des projets valables.

IV. Communiquer les résultats de la recherche

Au bout du compte, l'aide que l'on accorde à un programme dynamique d'analyse quantitative des questions sociales dépend de l'appui et de l'intérêt du grand public. De fait, c'est en exécutant des recherches pertinentes et en portant leurs résultats à l'attention des personnes qui peuvent s'en servir pour les débats et pour la prise de décisions que l'on accroît la demande relative à des recherches du même type. Dans les meilleurs des cas, cependant, le processus visant à recruter et à encourager les chercheurs universitaires ne valorise guère les aptitudes ni le tempérament nécessaires pour éveiller l'intérêt du public. À part quelques exceptions remarquables, les décideurs estiment que les sciences sociales d'aujourd'hui ou bien ne se rapportent pas directement à leurs problèmes, ou bien s'intéressent à des questions trop étroites. Quant aux chercheurs, ils répugnent souvent à entreprendre des travaux qui leur paraissent orientés, à l'excès, vers des problèmes pratiques à court terme. Les liens entre les chercheurs en sciences sociales et les médias sont également assez faibles, en dépit de l'énorme appétit de la population pour des statistiques sur l'éducation, l'emploi, la santé, l'aide sociale, le vieillissement et d'autres questions de société. Dans les deux cas, les difficultés sont suscitées tout particulièrement par la dispersion de l'information, l'inaccessibilité du langage et des différences dans les calendriers de travail.

Bien que l'on ait accompli des progrès dans certains domaines de recherche, il faut nous montrer plus dynamiques pour favoriser le dialogue entre les spécialistes des sciences sociales et les utilisateurs potentiels des résultats de leurs recherches. Ce dont nous avons besoin, c'est d'un

changement de culture, qui ne peut se produire qu'après une assez longue période d'interaction – et cette interaction n'aura lieu que si des occasions sont créées et exploitées systématiquement à cette fin. C'est pourquoi nous proposons l'établissement de forums de recherche, de même que la mise sur pied, au CRSH, d'un programme de communication en statistiques sociales.

A. Les Forums de recherche

> Nous proposons la mise sur pied d'un programme d'appui à des forums de recherche visant à mettre en contact des spécialistes de la recherche en sciences sociales, appartenant soit à des universités, soit à des administrations publiques ou à des organismes de recherche non gouvernementaux, avec des analystes de politiques et des décideurs de divers milieux. Ces forums parraineraient des conférences et des activités connexes qui fourniraient aux chercheurs des milieux gouvernementaux et universitaires l'occasion de discuter entre eux, et de profiter des recherches, des données et de l'expérience des uns et des autres.

Il arrive trop souvent, aujourd'hui, que les résultats de recherches universitaires sur des questions sociales ou économiques qui présentent un intérêt pour les décideurs de tous les paliers de gouvernement, fédéral, provincial et municipal, ne soient jamais communiqués à ces auditoires potentiels. Par ailleurs, les chercheurs universitaires ne sont pas au courant de nouveaux débats politiques sur lesquels leur expertise pourrait apporter un éclairage utile. En conséquence, on laisse inexploitée une extraordinaire réserve de connaissances qui favoriseraient une meilleure compréhension des tendances de la société canadienne et des enjeux politiques qui la concernent. Afin de remédier à cette situation, nous proposons un programme de forums de recherche.

Nos forums de recherche seraient créés sur le modèle du Forum canadien de recherche sur l'emploi (FCRE), qui a été établi en 1991 pour encourager la recherche au Canada et pour rassembler les spécialistes de la recherche axée sur les politiques en poste à Ottawa, d'une part, et les économistes du travail appartenant à des universités, d'autre part. Il y a dix ans à peine, les économistes du travail canadiens s'intéressant à la recherche inductive se servaient principalement de données américaines. Beaucoup de professeurs d'université et d'étudiants canadiens des cycles supérieurs en savaient davantage sur le fonctionnement et l'évaluation des systèmes d'aide sociale et d'assurance-chômage des États-Unis que sur les mécanismes de leurs équivalents canadiens. À très peu de frais, le FCRE a vite éliminé ce problème, et a créé un milieu dynamique et interactif de chercheurs provenant tant des universités que de la fonction publique. Pour de l'information complémentaire sur les activités du FCRE, prière de se reporter à l'**annexe E** (page 36).

Nous croyons que la formule du FCRE peut être adoptée dans d'autres domaines de recherche, et appliquée à un éventail élargi d'activités visant à réunir les chercheurs et les analystes de politiques d'une grande diversité d'établissements. Pour qu'un forum soit une réussite, il faut une masse critique de chercheurs travaillant dans un même domaine, bien défini sans être trop spécialisé. Les Forums auraient principalement pour objet d'organiser des conférences, en vue de communiquer les résultats de recherches aux décideurs et aux analystes de politiques, de permettre aux chercheurs de mieux comprendre les débats relatifs à l'orientation des politiques, et plus généralement d'offrir un cadre qui facilite les relations entre les différents secteurs intéressés par le Système canadien de recherche utilisant les statistiques sociales. Nous y voyons aussi une excellente occasion d'intégrer dans un milieu de recherche les étudiants diplômés et les jeunes chercheurs. Les Forums pourraient également fournir une tribune

commode et efficace où les chercheurs des universités comme de la fonction publique pourraient discuter des avantages et des inconvénients que présentent les enquêtes de Statistique Canada relativement à l'étude des questions politiques actuelles, et communiquer leurs observations à Statistique Canada.

Nous proposons que le CRSH tienne des concours, soumis au processus d'examen par les pairs, en vue de créer des forums de recherche quantitative en sciences sociales. Chaque forum serait financé pour une période de cinq ans, cette aide étant renouvelable aussi longtemps que l'on atteindrait les objectifs du programme. Nous envisageons la mise sur pied d'un ou deux forums par année pendant les cinq premières années. Par souci de simplicité, les demandes d'aide financière pourraient être soumises au même comité qui étudie les candidatures de groupes de recherche et de formation. Pour être admissible au programme, un groupe de chercheurs devrait former un comité de direction, composé d'une douzaine de membres qui seraient considérés comme codemandeurs. Il serait bon d'encourager la nomination, à ce comité de direction, non seulement de chercheurs universitaires, mais aussi de chercheurs de la fonction publique et d'organismes à but non lucratif. Il conviendrait aussi, dans l'évaluation des demandes, de considérer comme éléments positifs la participation d'autres organismes et, là où c'est possible, l'attribution d'une aide financière par ces derniers, notamment les universités, les ministères et les organismes non gouvernementaux qui s'occupent de questions d'intérêt public.

Les demandes se rapportant à de nouveaux forums devraient inclure un projet de conférence inaugurale pour la première année, ainsi qu'une indication de l'orientation des travaux pour les années suivantes. Il va de soi que les groupes existants dont les activités correspondent à ce que nous avons décrit plus haut pourraient solliciter un appui. L'aide financière vise principalement l'activité prédominante des Forums, soit l'organisation de conférences; des sommes seront prévues pour les tâches de secrétariat et le soutien logistique, pour les déplacements (tant des professeurs que des étudiants diplômés) et pour de modestes honoraires s'appliquant à la préparation de communications qui amènent le chercheur à s'éloigner significativement de son programme de recherche pour aborder de nouvelles questions. Afin de favoriser la diffusion des communications et la continuité des échanges, l'aide financière devrait couvrir le coût d'entretien d'un site Web; il faudrait accorder aussi un appui pour la publicité à donner aux conférences. De plus, comme la réussite d'un forum reposera en grande partie sur le travail, largement bénévole, des universitaires qui organiseront les conférences, assureront le développement de l'organisme et s'occuperont de son fonctionnement, on devrait faire en sorte de dégager d'une partie de leurs tâches d'enseignement les universitaires canadiens membres du comité de direction. Il faut également une aide financière pour les déplacements des membres qui se rendent aux réunions du comité.

B. Le Programme de communication en statistiques sociales

Nous recommandons que le CRSH mette sur pied un programme de communication en statistiques sociales, qui aurait expressément pour but de faire mieux connaître la recherche quantitative en sciences sociales auprès de l'ensemble de la population. En collaboration avec les Groupes de recherche et de formation, les Forums de recherche et les bénéficiaires de bourses de recherche de niveau supérieur, et en consultation avec les services de communication des universités et de Statistique Canada, ce programme viserait à mettre en valeur les résultats de recherches utilisant les statistiques sociales, de même qu'à favoriser les débats entourant la publication de telles statistiques.

Les statistiques se prêtent tout particulièrement bien à la communication en sciences sociales, en raison du flot constant d'information quantitative que publient les médias relativement aux phénomènes économiques et sociaux. Divers segments de la population, au même titre que les organismes non gouvernementaux, en sont venus à attendre ces chiffres, tout comme les débats qui accompagnent leur diffusion. Vu ces circonstances favorables, les statistiques sociales devraient devenir prioritaires dans la stratégie de communication du CRSH. Le Programme de communication en statistiques sociales que mettra sur pied cet organisme devrait être conçu en fonction de deux sources d'information, soit la recherche universitaire et les publications périodiques de Statistique Canada.

En premier lieu, beaucoup de résultats de recherches que produisent les universitaires présentent un immense intérêt pour le grand public et pour les décideurs, et ce sera encore davantage le cas avec la création des Groupes de recherche et de formation. Le paradoxe, cependant, c'est qu'il est avantageux de gérer les communications à partir d'un centre, ou du moins à partir de quelques endroits bien choisis où l'information peut aisément être mise à la disposition de tous ceux qui pourraient en faire usage, alors que la recherche financée par les programmes du CRSH est exécutée par une multitude de groupes et de centres de recherche, répartis dans un grand nombre d'établissements d'études supérieures. La difficulté consiste donc à réunir et à organiser l'information, de même qu'à préparer les chercheurs à la transmettre avec efficacité, tout en ayant rarement la possibilité d'influer directement sur la façon dont aura lieu la communication.

Le Programme de communication devrait tout d'abord surveiller le volume de la recherche utilisant les statistiques sociales, au moyen de ses liens avec les Groupes de recherche et de formation, avec les bénéficiaires de bourses de recherche de niveau supérieur en statistiques sociales, et surtout avec les Forums de recherche. De fait, le Programme devrait être représenté systématiquement au comité de direction des Forums de recherche. La plupart de ces forums auront sans doute une page Web, et il y aurait lieu de créer une page Web centrale qui donnerait accès à toutes les autres. Il conviendrait aussi d'appeler l'attention des médias sur les recherches intéressantes et pertinentes qui seront produites ou présentées par les Forums; un bulletin pourrait être publié régulièrement, où serait donné pour chaque article le nom de personnes avec qui communiquer au sujet des résultats. Les responsables du Programme devraient en outre avoir des liens suivis avec les services de communication des universités, afin de les inciter à mettre davantage en valeur la recherche utilisant les statistiques sociales. On pourrait arriver à ce résultat en invitant les universités à signaler ce type de recherche à la direction du Programme, et en faisant en sorte que le Programme accorde une aide aux services de communication de ces établissements, au besoin. Enfin, le Programme pourrait offrir aux chercheurs, à l'occasion par exemple du Congrès des sciences sociales, des ateliers de formation sur les relations avec les médias.

En deuxième lieu, Statistique Canada publie régulièrement une diversité de statistiques sociales et économiques, ce qui porte certaines questions à l'attention du grand public. Le CRSH pourrait tirer parti de l'intérêt que l'on suscite de cette manière pour mieux faire connaître les idées des universitaires spécialistes des sciences sociales auprès des médias et de la population en général. Les objectifs et le mode de fonctionnement du Programme de communication du CRSH resteront sans doute bien distincts de ceux de la Division des communications de Statistique Canada, mais il serait utile de favoriser la collaboration entre les deux services. Par exemple, si on lui envoyait d'avance le calendrier des publications de Statistique Canada, l'équipe du Programme de communication du CRSH pourrait faire parvenir aux médias la liste à jour des experts dans n'importe quel domaine de recherche, et suggérer les noms de chercheurs qui seraient bien placés pour participer à des débats sur les statistiques sociales. Étant donné le caractère décentralisé de nombreux aspects de la politique sociale, on pourrait s'efforcer de trouver des experts en provenance des diverses régions du Canada. L'équipe du Programme pourrait organiser des conférences et des débats entre experts, sur des questions susceptibles de bénéficier de l'éclairage apporté par les statistiques sociales. Ces activités seraient souvent conçues pour les médias, mais on pourrait viser aussi un auditoire de fonctionnaires, d'hommes

et de femmes politiques ou de représentants d'organismes divers. Il serait utile, à ce propos, d'obtenir la collaboration de la Fédération canadienne des sciences humaines et sociales, avec laquelle il serait possible de partager les tâches.

Enfin, l'équipe du Programme pourrait, de concert avec les Forums de recherche, organiser des ateliers et des tournées de conférences qui donneraient l'occasion de présenter à divers groupes, en particulier à de jeunes universitaires qui envisagent de se consacrer à la recherche en statistiques sociales, les résultats de travaux faisant appel à l'analyse quantitative. Cette activité pourrait être entreprise en collaboration avec Statistique Canada, qui est déjà actif dans ce domaine.

V. Coordination, coûts et plan d'exécution

A. La coordination du Système canadien d'analyse des statistiques sociales

> Afin de coordonner la mise en oeuvre du programme que nous préconisons et de contrôler l'avancement des travaux, nous recommandons que le CRSH et Statistique Canada concluent un protocole d'entente précisant les objectifs de l'initiative, à court et à long terme. Ils devraient aussi nommer conjointement un coordonnateur qui serait chargé de l'ensemble du Système de recherche, et un comité consultatif composé de cadres supérieurs des universités et d'autres organismes.

L'initiative que nous avons décrite en matière de recherche quantitative en sciences sociales comporte un ensemble de programmes interdépendants. Un milieu qui favorise l'excellence de la recherche suppose un certain nombre d'établissements qui se soutiennent mutuellement – c'est pourquoi nous parlons de la mise sur pied d'un *système* canadien d'analyse des statistiques sociales. La coordination continue des diverses composantes du Système pourrait être favorisée par un accord explicite sur les buts qu'on se propose et sur le mode d'organisation officiel des éléments du programme. Les acteurs qui tiendront les différents rôles que prévoit le Système auront besoin de recevoir d'une manière constante, les uns sur les autres, une plus grande quantité d'information que ne peuvent fournir des réseaux informels. Il y a aussi des questions primordiales qui ne relèvent d'aucune composante en particulier, comme les efforts à fournir pour faire progresser des perspectives intellectuelles et méthodologiques s'appliquant à la totalité des domaines et des activités de recherche. Nous recommandons que l'établissement des buts et les travaux de coordination fassent l'objet d'un protocole d'entente entre le CRSH et Statistique Canada, suivi de la création d'un poste de coordonnateur.

Nous recommandons la nomination d'un coordonnateur à plein temps, au niveau de directeur. La personne choisie devrait posséder une vaste expérience du courtage et de la gestion des travaux de recherche, de même que l'envergure intellectuelle et l'expérience de la recherche nécessaires pour communiquer efficacement avec un éventail de spécialistes des sciences sociales. Le mandat du coordonnateur devrait être en partie de créer et de développer des partenariats qui faciliteront l'entreprise de recherche dans son ensemble.

Le coordonnateur pourrait réunir un groupe de liaison dont feraient partie des représentants de toutes les composantes du Système : les présidents et les agents d'administration des comités d'examen par les pairs qui s'occupent des Groupes de recherche et de formation, des bourses et des Forums; des membres des Groupes de recherche et de formation et des Forums de recherche; les coordonnateurs du Programme d'été et des Centres d'accès aux données de

recherche, ainsi qu'un représentant du Programme de communication. Le coordonnateur relèverait du statisticien en chef et du président du Conseil de recherches en sciences humaines.

En vue de nouer et de préserver des liens solides avec le milieu universitaire et avec celui de la recherche axée sur les politiques, nous recommandons en outre la création d'un comité consultatif de niveau supérieur, dont feraient partie des dirigeants des universités (p. ex. des vice-recteurs responsables de la recherche) et de la fonction publique (p. ex. des sous-ministres adjoints responsables des politiques), ainsi que des chercheurs de l'étranger ou des directeurs de recherches en provenance de pays où l'on est parvenu à mettre en oeuvre des programmes similaires.

B. Coûts et plan d'exécution

L'initiative que nous proposons réunit un bon nombre de composantes étroitement liées. Le Groupe de travail mixte a établi provisoirement le coût de ces composantes en se fondant sur celui de programmes du même genre. Nous reconnaissons toutefois qu'il faudrait effectuer un calcul beaucoup plus complet à mesure que les propositions se transformeraient en plans de travail détaillés.

Nous reconnaissons également qu'il faudra un certain nombre d'années pour que soit mise en oeuvre la totalité du programme, et nous pensons qu'il faudrait en échelonner la réalisation sur une période de cinq ans, à compter de l'exercice 1999-2000. Nous estimons qu'il convient de procéder d'urgence à l'établissement d'un certain nombre de composantes, et que les travaux pourraient et devraient commencer au cours du prochain exercice. La priorité devrait être accordée à la mise sur pied, au moins comme projets pilotes, de plusieurs centres d'accès aux données de recherche. Il devrait être également possible de lancer le Programme de formation au moyen de quelques cours bien choisis, et de constituer quelques forums de recherche en organisant un certain nombre de conférences sur des sujets prioritaires. Il conviendrait aussi d'entreprendre l'élaboration, au CRSH, d'une stratégie de communication relative aux statistiques sociales. Nous nous rendons compte que les programmes d'envergure donnant lieu à des concours, en particulier les Groupes de recherche et de formation et les Bourses de recherche de niveau supérieur, exigeront un processus d'examen relativement long; il faudrait donc les faire démarrer le plus tôt possible.

Annexe A

Certains domaines prioritaires pour la recherche en statistiques sociales

Divers enjeux politiques et domaines de recherche profiteraient d'une intensification de la recherche quantitative à partir des bases de données disponibles. Le rôle du Groupe de travail mixte ne consiste toutefois pas à choisir uniquement des « options gagnantes », mais plutôt à établir des programmes comportant des processus de sélection équitables. Les brèves descriptions qui suivent visent à suggérer des domaines de recherche que cette initiative pourrait stimuler, sans toutefois se limiter à ces domaines.

1. Développement de l'enfant

Au cours des dix dernières années, on est devenu de plus en plus conscient des effets à long terme de la qualité des expériences des enfants au cours de leurs premières années de vie sur leur bien-être et leur bonheur, leurs expériences futures en matière scolaire et professionnelle, ainsi que leur état de santé comme adultes.

Un programme de recherche sur le développement des jeunes enfants au Canada pourrait être axé initialement sur les questions suivantes.

- Quelle est la proportion d'enfants canadiens qui sont exposés à des expériences de vie excessivement négatives, du fait de la pauvreté, de la violence familiale, des compétences inappropriées des parents et des préjugés raciaux ou ethniques?
- Dans quelle mesure la vulnérabilité de l'enfant est-elle liée à la structure familiale, et plus particulièrement dans le cas des familles monoparentales, par opposition aux familles biparentales, ainsi qu'aux facteurs socio-économiques comme le revenu familial et le niveau de scolarité des parents?
- Dans quelle mesure les réponses à ces deux premières questions dépendent-elles de l'âge de l'enfant et de la cohorte d'âge?
- Quels sont les mécanismes tampons ou les facteurs de protection liés à un développement sain de l'enfant?

Au fur et à mesure que les intervenants, y compris les parents, les enseignants, les administrateurs et les décideurs gouvernementaux, tentent d'élaborer une nouvelle politique sociale pour le Canada, ils doivent trouver des façons de soutenir les familles et les collectivités, sans que cela entraîne une augmentation trop importante des dépenses gouvernementales. Chez les praticiens, il est courant de penser que les interventions cliniques pour tous les enfants à risque sont trop coûteuses et ne répondent pas de façon appropriée aux besoins de la totalité d'entre eux. Toutefois, les interventions visant des groupes particuliers constituent-elles une meilleure option, ou des programmes universels sont-ils plus susceptibles d'avoir des répercussions plus grandes? Le cas échéant, quels types de programmes seraient les plus efficaces? La recherche sur le développement de l'enfant pourrait fournir un moyen d'évaluer les progrès accomplis en vue de réduire la vulnérabilité des enfants et de redresser les inégalités.

Pour résoudre ces questions, la principale source de données est l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ).

2. Jeunes en transition

La transition des études secondaires aux études postsecondaires, ainsi que des études au marché du travail, pose un problème pour de nombreux jeunes. Même si les documents de recherche font état des facteurs familiaux qui sont principalement liés aux faibles résultats

scolaires et à la réussite professionnelle, une somme considérable de recherche est nécessaire pour comprendre les cheminements qui mènent à la réussite et les transitions critiques que doivent effectuer les jeunes âgés de 15 à 25 ans. Un programme de recherche devrait être guidé par certaines des grandes questions qui suivent :

- Quels sont les groupes les plus susceptibles d'abandonner l'école avant d'avoir obtenu leur diplôme et ceux qui ont le plus de difficulté à faire la transition vers les études postsecondaires ou le marché du travail?
- Quels sont les compétences, les aptitudes et les comportements des jeunes qui effectuent avec succès ces transitions et de ceux qui échouent?
- Dans quelle mesure les taux d'obtention des diplômes varient-ils entre les écoles dans chaque province? Quels sont les facteurs scolaires qui contribuent à des taux élevés d'obtention des diplômes et à de bons résultats scolaires?
- Des programmes comme l'alternance travail-études, le mentorat, la formation à distance et les stages aident-ils les étudiants à effectuer ces transitions?

Au moins cinq ensembles de données nationales et internationales peuvent s'appliquer à ces questions : les données de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), du Programme international pour le suivi des élèves, de l'Enquête planifiée auprès des jeunes en transition au Canada, de même que celles sur l'alphabétisation au niveau international, ainsi que des enquêtes sur la formation et l'éducation.

3. Familles en évolution

Au cours des trois dernières décennies, la vie familiale s'est modifiée considérablement. La proportion de plus en plus importante de femmes sur le marché du travail (particulièrement de mères ayant de jeunes enfants), ainsi que la croissance des régimes de travail souples et des emplois atypiques, ont modifié la répartition du travail entre les sexes, tant à l'intérieur des familles que dans la société, et ont mené à une réorganisation de l'emploi du temps des familles. L'augmentation du nombre de séparations et de divorces, le déclin du mariage et l'augmentation du nombre de cohabitations ont transformé les trajectoires des femmes, des hommes et des enfants dans la famille. Les chercheurs commencent à peine à comprendre les répercussions considérables de ces changements. Parmi les questions auxquelles il faudrait répondre figurent les suivantes :

- Quels sont les effets de l'évolution du monde du travail sur la propension des hommes et des femmes à établir et à maintenir des relations conjugales et familiales?
- Comment les rapports existants entre les membres de la famille (p. ex. entre les conjoints, entre les parents et les enfants, entre les enfants et les grand-parents) se modifient-ils et sont-ils redéfinis une fois que les parents se séparent?
- Quelles sont les conséquences de la rupture de la famille sur la paternité?
- Comment les ressources sont-elles partagées entre les conjoints une fois qu'une union prend fin?

On a besoin de données longitudinales pour répondre à la majorité de ces questions. Ces ensembles de données comprennent celles de l'Enquête sur la famille de 1984, des Enquêtes sociales générales de 1990 et de 1995 sur la famille, de l'Enquête longitudinale nationale sur les enfants et les jeunes et de l'Enquête sur la dynamique du travail et du revenu.

4. Vieillir au Canada

Ces dernières années, les chercheurs, les décideurs et les gestionnaires de programme ont commencé à s'intéresser au vieillissement de la population et à la situation des personnes âgées au Canada. Parmi les principales lacunes que comportent nos connaissances sur le vieillissement au Canada figurent les suivantes :

- Des descriptions précises des inégalités qui entourent le processus de mise à la retraite des Canadiens, à partir de l'âge de 50 ans. Les politiques et la recherche au Canada reposent pour une large part sur l'hypothèse erronée que toutes les personnes entrent sur le marché du travail vers l'âge de 20 ans et prennent leur retraite vers l'âge de 65 ans.
- Des connaissances au sujet des personnes les plus âgées, c'est-à-dire celles de 85 ans et plus. Des questions importantes ont trait à la santé de ce groupe, aux soins qui lui sont donnés, à ses conditions de logement et à la sécurité du revenu des personnes qui bénéficient de prestations de retraite limitées (y compris le RPC/RRQ).
- Les caractéristiques des personnes qui se retrouveront probablement dans des foyers pour personnes âgées ou qui y résident actuellement. Les connaissances concernant ces groupes sont essentielles pour la planification et les politiques en matière de santé, de logement et de sécurité économique.

Il existe de nombreux ensembles de données pour répondre à ces questions, y compris celles de l'Enquête nationale sur la santé de la population, de l'Enquête sur la dynamique du travail et du revenu, des cycles de l'Enquête sociale générale portant sur des sujets connexes, du Réseau canadien de recherche sur le vieillissement, des fichiers administratifs longitudinaux, de l'Étude sur la santé et le vieillissement au Canada, de l'Enquête sur le vieillissement et l'autonomie, de l'Enquête sur la santé et la limitation d'activités et de l'Enquête sur les établissements de soins pour bénéficiaires internes.

5. Éducation, compétences et alphabétisation

Le développement des compétences et du capital humain, ainsi que les progrès technologiques (« logiciels » et « équipement »), sont considérés comme les principaux éléments qui sous-tendent la productivité, et par conséquent, le niveau de vie dans les économies modernes. Il est difficile de surestimer l'importance que revêtent l'éducation, la formation et la mise en valeur des compétences pour la plupart des sociétés. Même si ce sujet a été au centre d'efforts de recherche majeurs, les changements rapides qui touchent l'économie et la société, et plus particulièrement ceux qui ont trait au rôle et à l'importance de la scolarité et des compétences, doivent faire l'objet de recherches permanentes. Dans une société « axée sur le savoir », les questions qui suivent figurent déjà dans le programme d'action.

- Le niveau de réussite des étudiants canadiens dans un contexte national et international.
- Les effets des changements qui touchent le système d'éducation sur l'accès à l'enseignement supérieur.
- L'apprentissage toute la vie durant et ses répercussions.
- Le rôle de l'alphabétisation, *indépendamment* de la réussite scolaire, à l'égard de la réussite sur le marché du travail et des activités quotidiennes.
- La pertinence de la formation assurée par les entreprises canadiennes.
- Le lien entre le capital humain et les changements technologiques.
- Les pénuries de compétences et les surplus.
- Le rôle du capital humain dans l'accroissement de la richesse.

Un bon nombre de sources de données existantes portent sur ces questions, y compris, l'Enquête sur l'éducation et sur la formation des adultes, l'Enquête internationale sur l'alphabétisation des adultes, les tests d'évaluation du rendement scolaire, l'Enquête de suivi auprès des diplômés,

l'Enquête planifiée sur le lieu de travail et les employés, ainsi que des enquêtes traditionnelles, comme l'Enquête sur la population active et les recensements.

6. Répartition des salaires et du travail

Les économies développées d'aujourd'hui font face à des enjeux politiques essentiels touchant les changements quant à la répartition des salaires et du travail. De façon plus particulière, de nombreux pays, y compris le Canada, les États-Unis et la Grande-Bretagne, ont connu une augmentation substantielle des inégalités salariales au cours des vingt dernières années. Parmi les questions clés du point de vue de la recherche figurent les suivantes :

- Qu'est-ce qui a été à la source de ces tendances : des importations accrues à partir de pays où la rémunération est faible, ainsi que le phénomène connexe de la sous-traitance confiée à ces pays, ou encore l'avènement de nouvelles technologies qui éliminent les emplois des travailleurs les moins qualifiés?
- Le déclin des salaires des hommes non qualifiés est-il le résultat de la diminution de l'influence des syndicats? Les changements qui ont touché la qualité de notre système d'éducation ont-ils joué un rôle? L'offre relative de travailleurs très scolarisés influence-t-elle l'inégalité? Si plus d'un de ces facteurs entrent en jeu, quelle est leur importance relative et les relations qui existent entre eux?
- La structure institutionnelle d'un pays a-t-elle des répercussions sur le problème du chômage (comme en France et en Allemagne), ou le problème de l'inégalité salariale (comme aux États-Unis), ou les deux (comme au Canada)? La politique relative au marché du travail donne-t-elle suite aux conclusions des recherches et, le cas échéant, comment?
- Quel rôle les systèmes de soutien social joueront-ils à l'avenir?

Des recherches approfondies ont été effectuées sur ces sujets, mais de nombreuses questions subsistent. De nouvelles sources de données, par exemple, l'Enquête sur la dynamique du travail et du revenu et l'Enquête sur le lieu de travail et les employés, ainsi que des sources plus traditionnelles, comme l'Enquête sur la population active, qui se tient sur une base mensuelle, jetteront un nouvel éclairage sur le rôle des changements technologiques axés sur les compétences et les questions connexes.

7. Soutien social et communautaire

Au cours de la dernière décennie, on a commencé à reconnaître davantage l'importance des activités non rémunérées, y compris les soins aux enfants, les travaux ménagers, les soins aux personnes âgées et le bénévolat. Évidemment, les femmes continuent d'effectuer la majorité du travail non rémunéré, en dépit de leur participation plus grande à la population active. Dans le cadre du Projet de recherche sur les politiques du gouvernement fédéral, on a déterminé que la modification de l'emploi du temps tout au long de la vie et à chaque étape de celle-ci est à la base de nombreux défis qui se posent aujourd'hui sur le plan des politiques sociales. Les soins aux enfants et aux personnes âgées font aussi l'objet de changements. Le vieillissement de la population canadienne ainsi que l'évolution des responsabilités relatives aux soins des personnes, celles-ci étant transférées des établissements aux personnes et aux familles, constituent des défis majeurs. Les résultats combinés de la baisse de l'âge de la retraite et de l'augmentation de l'espérance de vie peuvent aussi donner lieu à des déséquilibres du point de vue de l'emploi du temps pour les personnes plus âgées. Des questions se posent de plus en plus au sujet de l'érosion du soutien communautaire et du « capital social ».

Parmi les données qui permettent de résoudre ces questions figurent : les documents relatifs aux diverses années de l'Enquête sociale générale portant sur l'emploi du temps ainsi que sur le soutien social et les soins aux autres, l'Enquête nationale sur le don, le bénévolat et la participation, les recensements nationaux, l'Enquête nationale sur la santé de la population,

l'Étude sur la santé et le vieillissement au Canada et l'Enquête longitudinale nationale sur les enfants et les jeunes.

8. Répercussions sociales des sciences et de la technologie sur les familles/ les enfants et sur le bien-être

Les sciences et la technologie sont des forces dominantes du présent siècle, et peut-être même *les* forces dominantes. On prétend souvent qu'elles amélioreront la qualité de vie et les perspectives d'avenir des enfants, mais nous savons peu de choses au sujet des répercussions sociales à long terme des changements technologiques. Il est difficile de tirer des conclusions générales quant au caractère positif de ces répercussions ou, de façon générale, de décrire comment elles fonctionnent. Par ailleurs, il existe des aspects mal compris de la répartition. Qui profite des changements et qui en pâtit? Par exemple, quels sont les effets à long terme de l'écart entre les enfants qui ont accès à un ordinateur et à Internet, à l'école et à la maison, et ceux qui n'y ont pas accès?

On sait peu de choses au sujet des répercussions sociales des sciences et de la technologie sur les familles et les enfants, ainsi que sur le bien-être. Les efforts déployés actuellement par Statistique Canada pour établir une structure cohérente, en vue de l'élaboration systématique de données statistiques sur les sciences et la technologie, ouvrent des perspectives du point de vue de l'exploitation analytique des données existantes et de la production de nouveaux outils et couplages de données.

Un bon nombre d'ensembles de données existants, et d'autres qui seront disponibles sous peu, peuvent contribuer à résoudre ces questions, par exemple, les données de l'Enquête sur les innovations, de l'Enquête sur le lieu de travail et les employés, de l'Enquête de suivi auprès des diplômés, des enquêtes spéciales sur l'utilisation d'Internet, les capacités de recherche, etc. Il faut aussi mettre au point de nouvelles données pour élargir les domaines de recherche.

9. Évolution du milieu de travail et utilisation de la technologie

Les années 1990 ont vu converger un certain nombre de phénomènes liés à la technologie, qui ont des répercussions sur le lieu de travail et les travailleurs, notamment les suivants :

- L'utilisation accrue des technologies de l'information, y compris la montée d'Internet et des technologies connexes de communication, dans presque tous les secteurs d'activité, le tout s'accompagnant de préoccupations quant à l'obtention des gains de productivité souhaités.
- L'accent accru qui est mis sur l'importance de l'innovation pour la survie et la croissance des entreprises et pour l'augmentation de la productivité.
- Les répercussions de la technologie sur la réduction des effectifs.
- Des préoccupations au sujet de la plus grande polarisation de la société que pourrait entraîner la technologie.
- L'accent qui est mis sur les questions relatives aux ressources humaines, comme la formation, les méthodes de rémunération, les horaires de travail et les nouvelles pratiques mises en œuvre en milieu de travail pour atteindre un niveau de rendement élevé.

On sait relativement peu de choses sur l'adoption et la diffusion de la technologie et de l'innovation dans les divers milieux de travail ainsi que leurs répercussions sur le lieu de travail et les travailleurs. Même si les chercheurs se préoccupent depuis longtemps de ces questions, il existe peu de sources de données à grande échelle pour documenter le taux d'utilisation de la technologie et des innovations, sans parler de leurs répercussions. Cela a mené à l'utilisation

d'approximations souvent contestables ou d'études de cas, en ce qui a trait à l'utilisation de la technologie.

Plus récemment, de nouveaux ensembles de données ont été constitués pour nous aider à mieux comprendre les questions dans ce domaine et pour ouvrir de nouveaux débouchés à la recherche. On pense notamment aux enquêtes sur la technologie, à celles sur l'innovation, à l'Enquête sur le lieu de travail et les employés, ainsi qu'à une enquête sur les facteurs déterminants de la croissance des entreprises.

10. Bien-être, revenu et pauvreté

Les recherches sur le bien-être et les inégalités matérielles portent sur trois questions principales. Tout d'abord, quelle est la répartition du revenu, de la richesse et des besoins essentiels en matière de logement et de nourriture entre les personnes, les familles, les ménages et les collectivités? En deuxième lieu, quelles sont les conséquences de ces inégalités sur la qualité de la vie sociale des personnes, de façon générale et pour des aspects particuliers comme la santé? Enfin, quels sont les mécanismes qui font en sorte que les inégalités se reproduisent ou se modifient au fil des ans, au cours de la vie d'une personne et d'une génération à l'autre? Sur un plan global, cette question peut être reformulée du point de vue de l'évolution des collectivités géographiques et d'autres formes de collectivités au fil des ans.

Des études sur le bien-être, le revenu et la pauvreté sont nécessaires pour comprendre les effets d'une très grande diversité de politiques comportant des dépenses considérables, par exemple : les effets redistributifs de l'impôt sur le revenu, l'efficacité des programmes de bien-être social, d'éducation et de santé, ainsi que les modalités de prestation des services sociaux et des soins de santé.

On dispose depuis plusieurs décennies de profils de base des inégalités économiques et sociales. Toutefois, au fur et à mesure que les politiques et les programmes évoluent, de nouvelles recherches sont nécessaires pour évaluer leurs répercussions. Les nouvelles enquêtes longitudinales fournissent donc une façon nouvelle d'examiner de près le processus de changement au fil des ans. Il est possible d'évaluer les répercussions sur le bien-être des Canadiens de changements comme la perte d'un emploi ou la fin d'une union conjugale. On peut aussi évaluer les répercussions à court et à long termes des privations graves et persistantes sur le bien-être des enfants et des jeunes, ainsi que sur leur santé mentale et physique.

Parmi les données disponibles pour répondre à ces questions figurent celles de l'Enquête sur les finances des consommateurs, de l'Enquête sur la dynamique du travail et du revenu, de l'Enquête longitudinale nationale sur les enfants et les jeunes, de l'Enquête nationale sur la santé de la population, de l'Enquête sur les dépenses des familles et des recensements.

Annexe B

Aperçu d'un programme de cours d'été

Même si plusieurs années seront peut-être nécessaires pour atteindre la stabilité, il est utile de penser à la forme que pourrait prendre un programme de cours d'été plus élaboré, qui serait réparti sur quatre semaines. En ce qui a trait à l'analyse des données, le programme reposerait sur deux cours, chacun d'une durée de deux semaines. Le premier servirait d'introduction de base à la régression et serait destiné aux participants qui ont peu de connaissances dans ce domaine. Il reposerait sur des exemples d'analyse de données tirées d'enquêtes transversales, comme l'Enquête sociale générale. Un deuxième cours de deux semaines, sur les modèles linéaires, élargirait les techniques de régression aux données tronquées, aux données de type nominal, aux dénombrements et aux données de durée. Les deux cours combineraient des présentations le matin et des exercices pratiques supervisés l'après-midi. Dans le cadre de ces cours et d'autres, on devrait mettre l'accent sur la façon de rendre compte des résultats de l'analyse de données et assurer une formation précise à ce sujet.

Outre les deux cours de base sur l'analyse des données, chaque programme de cours d'été devrait inclure, en alternance et en fonction de la demande, un ou plusieurs cours de niveau intermédiaire sur des sujets comme l'analyse longitudinale, les modèles hiérarchiques, les modèles de variables latentes et les données de type nominal. L'accent devrait être mis sur l'établissement d'un bassin de chercheurs familiers avec les techniques modernes d'analyse des données, en vue de procéder à des recherches à la fine pointe des connaissances. Il sera souvent approprié de mettre l'accent sur une enquête particulière, spécialement pour l'analyse longitudinale, étant donné que des efforts considérables doivent être déployés pour se familiariser avec un ensemble de données. On pourrait offrir une combinaison de cours d'une semaine et de deux semaines, selon le sujet. Dans le cadre d'ateliers plus restreints, on pourrait aborder, en alternance, des sujets d'analyse plus limités, par exemple, la pondération et l'estimation pour des échantillons complexes, l'analyse de données transversales regroupées et les méthodes statistiques robustes.

Deux statisticiens (un de Statistique Canada et un autre d'une université) ont proposé au Groupe de travail une troisième forme de cours de méthodologie, qu'il convient assurément de mettre à l'essai. Les groupes de formation mettraient l'accent directement sur un ensemble de données choisies selon les intérêts des enseignants et des étudiants. Les étudiants travailleraient en équipe, sous la direction d'un chercheur et avec l'aide d'un programmeur, d'un spécialiste des méthodes d'enquête, d'un spécialiste du sujet visé par l'enquête et d'un statisticien. Le cours serait axé sur l'utilisation des données choisies pour résoudre un problème de recherche particulier, les questions de conception, de statistiques et de justification étant abordées au fur et à mesure qu'elles se présentent.

On devrait envisager la possibilité de tenir un atelier d'une durée d'une semaine sur la philosophie, la logique et la stratégie de l'analyse des données, par exemple. Ces dernières années, on s'est davantage intéressé à expliciter les principes sous-jacents à l'analyse quantitative des données sociales et, par conséquent, à en débattre. Par exemple, on a examiné en long et en large les rôles relatifs de l'interprétation causale et de la description des phénomènes sociaux, par suite de la publication de l'ouvrage « Making It Count » de Stanley Lieberson. Dans le cadre d'un tel cours, on pourrait aborder la question suivante : dans quelle mesure et de quelle façon des données particulières et l'analyse des données peuvent nous informer sur les processus sociaux? On pourrait en même temps discuter des stratégies d'analyse, en mettant un accent particulier sur les ensembles de données d'enquêtes complexes et sur la façon d'effectuer une analyse pour rendre justice à un sujet complexe.

Outre les cours de statistiques appliquées, chaque séance d'été devrait inclure au moins un cours axé sur l'analyse d'une enquête, par exemple, l'Enquête sur la dynamique du travail et du revenu, l'Enquête sociale générale, pour un aspect particulier, l'Enquête nationale sur la santé de la population, ou encore l'Enquête longitudinale nationale sur les enfants et les jeunes. Ces cours pourraient être ouverts aux chercheurs prêts à s'engager à effectuer des recherches sur l'ensemble de données pertinent, et devraient être dirigés par deux chercheurs ou plus responsables de l'ensemble de données. Ils comprendraient en outre des dispositions explicites de mentorat pour les chercheurs débutants. Des chercheurs plus expérimentés pourraient suivre le cours principalement pour se familiariser avec un ensemble de données et pourraient collaborer au mentorat. Selon le sujet, la familiarisation aux données pourrait être combinée à un atelier statistique portant sur les méthodes nécessaires pour analyser des données. Au besoin, on pourrait tenir des activités de suivi, y compris une stratégie en vue de fournir des conseils méthodologiques sur une base permanente, des plans pour la révision des manuscrits et, si les ressources le permettaient, une réunion subséquente pour discuter de résultats de la recherche. Selon le sujet et l'accent mis sur la formation dans le cadre du cours, ce dernier pourrait durer d'une à deux semaines.

Chaque séance d'été devrait comprendre au moins un cours axé sur un sujet de recherche exhaustif, qui pourrait prendre la forme d'un séminaire. L'idée serait de réunir des chercheurs pour examiner des travaux dans un domaine donné et discuter des priorités de recherche. Cet exercice pourrait être couplé à un examen approfondi des sources de données disponibles sur le sujet. Il serait souhaitable que travaillent ensemble des chercheurs plus ou moins expérimentés et d'autres plus expérimentés, en vue de stimuler la recherche dans le domaine choisi, et les organisateurs devraient mettre au point une stratégie pour encourager les communications entre les chercheurs.

En alternance, le programme d'été devrait offrir des cours sur une variété de sujets. Parmi les idées intéressantes figure un séminaire sur la recherche sur les politiques, à partir de bases de données quantitatives. On pourrait faire appel à des chercheurs des ministères gouvernementaux, ainsi que d'organisations chargées des politiques, pour qu'ils décrivent leurs travaux ainsi que les préoccupations actuelles touchant la politique et la recherche dans leur secteur. On pourrait envisager de donner un petit cours sur les concepts de données. Il serait ainsi possible de mettre l'accent sur la façon dont Statistique Canada et les chercheurs d'enquête en général conceptualisent et mesurent les paramètres des variables sociales clés. Parmi les sujets qui pourraient être abordés figurent la race et l'origine ethnique, la participation au marché du travail, l'expérience de travail et les perceptions quant à la qualité de l'environnement social.

Étant donné qu'il ne semble pas exister de programmes des cycles supérieurs au Canada comportant des cours de formation avancée sur la conception d'enquêtes, on devrait envisager la possibilité d'inclure de tels cours dans le programme d'été. La formation relative à l'analyse des données devrait permettre de disposer de chercheurs qui connaissent les limites de leurs données, mais elle ne peut remplacer les connaissances comprises dans les documents spécialisés qui existent sur la conception d'enquêtes. Idéalement, le cours devrait commencer par un examen des recherches effectuées sur la conception des questionnaires et des stratégies d'essai préliminaire, d'évaluation et de révision de nouvelles enquêtes. Le tout serait suivi par un atelier dans lequel les participants rédigeraient, mettraient à l'essai et réviseraient un nouveau questionnaire. Idéalement, le deuxième volet du cours devrait être donné de concert avec Statistique Canada ou un organisme universitaire de recherche d'enquête.

Annexe C

Centres d'accès aux données de recherche

Afin de faciliter l'accès aux microdonnées confidentielles aux fins de la recherche, les chercheurs concernés seraient « réputés » être des employés de Statistique Canada. De nombreuses questions juridiques et organisationnelles doivent être résolues si l'on adopte ce scénario, et la présente annexe fait état d'une approche à cet égard. Parmi les questions à résoudre figure la proximité physique des données. L'accès sera difficile pour de nombreuses personnes si les données ne sont disponibles qu'à Ottawa. Il est donc proposé que l'approche dont il est question ici soit élaborée à l'intérieur d'une structure de Centres d'accès aux données de recherche.

1. Qu'est-ce qu'un Centre d'accès aux données de recherche?

Un Centre d'accès aux données de recherche est un emplacement physique qui constitue un prolongement de Statistique Canada et qui est doté d'un environnement sécuritaire permettant de protéger des fichiers de microdonnées confidentiels. Le centre compterait un programme de recherche affilié, qui serait administré par le CRSH et Statistique Canada. Les chercheurs, qui devraient être assermentés en vertu de la *Loi sur la statistique*, auraient accès à des fichiers de microdonnées confidentiels conservés dans le centre, ce qui leur permettrait de poursuivre leurs recherches. Ils seraient chargés d'assurer la confidentialité des données exactement de la même façon que les employés de Statistique Canada.

2. Enjeux majeurs liés à la mise en place d'un programme de Centres d'accès aux données de recherche

Protéger la confidentialité. Cet élément joue un rôle central, étant donné que la confidentialité constitue la pierre angulaire du système statistique. Tout accès aux données par les chercheurs qui ne sont pas des employés de Statistique Canada devra être conforme à la *Loi sur la statistique*.

S'assurer que les travaux et les chercheurs sont assujettis à la *Loi sur la statistique*. Afin d'assurer la confidentialité, les chercheurs doivent être assermentés en vertu de la *Loi sur la statistique* et être assujettis aux modalités de la Loi de la même façon que les employés permanents de Statistique Canada. La Loi prévoit des modalités d'accès aux données confidentielles, ces modalités devant être intégrées dans tout programme des Centres d'accès aux données de recherche.

S'assurer que les centres atteindront leur objectif. Les centres doivent être structurés de façon à pouvoir attirer des chercheurs de premier ordre et leurs étudiants, afin de veiller à ce que les objectifs énoncés précédemment dans le présent rapport soient atteints.

S'assurer que l'on dispose d'une structure d'organisation appropriée. Il existe des enjeux relativement à la façon dont les Centres d'accès aux données seront organisés et gérés ainsi qu'à la composition des comités d'approbation et d'évaluation des propositions de recherche. Il est proposé que le CRSH et Statistique Canada administrent conjointement le programme de recherche affilié, et que Statistique Canada gère les installations, à titre de prolongement de ses activités. Sur le plan opérationnel, la sélection et l'examen approfondi des programmes de recherche seraient effectués par des comités constitués de chercheurs réputés, en conformité des exigences de la *Loi sur la statistique* et de celles relatives à la confidentialité.

Assurer la protection matérielle appropriée des données. Afin de répondre aux exigences relatives à la confidentialité et de démontrer que ces exigences sont respectées, des dispositifs appropriés en vue d'assurer la sécurité matérielle des données doivent évidemment être mis en place.

Ces questions sont abordées dans l'aperçu qui figure ci-après.

3. Objectifs du programme de Centres d'accès aux données de recherche

Le programme comporterait deux objectifs :

- promouvoir la recherche quantitative effectuée par des chercheurs universitaires, des chercheurs d'organismes gouvernementaux, d'instituts de recherche ou d'autres organismes publics, à partir des fichiers de microdonnées de Statistique Canada, particulièrement ceux concernant les ménages;
- améliorer les programmes statistiques de Statistique Canada, grâce à la rétroaction assurée par les chercheurs utilisant les microdonnées, ainsi qu'aux recherches et aux documents produits par les chercheurs.

4. Éléments du programme de Centres d'accès aux données de recherche

Voici un aperçu des éléments essentiels d'un programme de Centres d'accès aux données de recherche.

- Un environnement informatique sécuritaire et convivial, permettant l'entreposage de microdonnées confidentielles. La sécurité devrait correspondre aux normes établies par Statistique Canada.
- Un processus de sélection, d'approbation et d'examen approfondi pour les recherches effectuées à partir des microdonnées confidentielles. Un comité composé de chercheurs d'expérience (universitaires et autres) dans un certain nombre de domaines, ainsi que de représentants du CRSH et de Statistique Canada, serait constitué pour administrer le processus de sélection et d'examen approfondi. Ce sont les chercheurs universitaires qui approuveraient les projets dans une très large mesure. Un comité d'examen par les pairs serait établi pour chaque grand domaine de recherche, par exemple, l'économie, la sociologie, la santé, l'éducation et la méthodologie statistique[1]. Les universitaires ou les chercheurs d'instituts de recherche, d'organismes gouvernementaux et d'autres organismes de recherche pourraient soumettre des propositions.
- Les Centres de données disposeraient d'au moins un employé de Statistique Canada en permanence, en vue d'en assurer la gestion matérielle, de fournir du soutien et de superviser les questions relatives à la confidentialité.
- Seules les personnes assermentées en vertu de la *Loi sur la statistique*, dans le cadre du programme de Centres d'accès aux données de recherche, auraient accès aux données des centres. Les chercheurs devraient signer un contrat faisant état de leurs obligations en vertu de la Loi[2].

---

[1] La préférence pourrait être accordée aux recherches empiriques qui sont pertinentes pour ce qui est des politiques. Afin de permettre au système de donner suite rapidement aux nouvelles questions politiques, on pourrait envisager un processus d'approbation à deux volets. Un volet servirait à prendre des décisions rapidement au sujet des études de faisabilité à court terme ne nécessitant que peu de financement, et un autre aurait trait aux projets à plus long terme, comportant des répercussions plus importantes au niveau des ressources.

[2] Afin de promouvoir la formation d'une nouvelle génération de chercheurs ayant l'expérience de l'utilisation de ces ensembles de données, ainsi que pour faciliter la recherche proprement dite, cela devrait s'appliquer aux assistants de recherche, particulièrement les étudiants des cycles supérieurs qui sont recrutés par un chercheur universitaire ou qui travaillent à une thèse sous sa direction.

- Les Centres devraient s'autofinancer, le financement provenant d'institutions responsables de leur gestion, de conseils de subvention comme le CRSH, et de chercheurs.

5. Fonctionnement des Centres en vertu de la *Loi sur la statistique*

Les chercheurs ayant accès à des données confidentielles doivent être assermentés en vertu de la *Loi sur la statistique*. Outre les employés permanents de Statistique Canada, la *Loi sur la statistique* restreint l'accès aux données confidentielles aux personnes « engagées à contrat pour des services spéciaux au ministre » (c'est-à-dire Statistique Canada). Ces personnes sont réputées être des employés de Statistique Canada. Le Bureau retient les services de chercheurs pour effectuer des travaux lorsqu'il ne dispose pas lui-même des ressources nécessaires. Statistique Canada n'a évidemment pas de ressources suffisantes pour produire les travaux d'analyse nécessaires pour exploiter les nombreux ensembles de données nouveaux (et plus anciens). Une telle exploitation est nécessaire pour contribuer à l'analyse de la valeur des progrès et des débats concernant les politiques gouvernementales, ainsi que pour la promotion de travaux de recherche fondamentale. Une part importante de la capacité nécessaire pour tirer parti de la richesse des sources de données réside à la fois chez les chercheurs universitaires et ceux d'autres organismes de recherche (autres organismes fédéraux, groupes de réflexion, gouvernements provinciaux, etc.).

Dans ce contexte et en vertu de la loi actuelle, pour qu'un chercheur puisse être assermenté, ses recherches doivent être similaires à celles que Statistique Canada effectuerait normalement (si des ressources étaient disponibles). Même si cela peut sembler restrictif, Statistique Canada effectue dans les faits une variété de travaux de recherche. Par conséquent, il ne devrait pas y avoir de problème. Dans le cadre du processus de sélection, les avantages d'une proposition donnée seraient prépondérants. Le comité de sélection et d'examen examinerait le processus d'examen une fois le projet terminé. Il est évident que Statistique Canada aurait son mot à dire dans le processus, étant donné qu'en vertu de la *Loi sur la statistique*, les travaux correspondent à des services spéciaux au ministre.

Obtention de l'approbation des répondants en vue du partage des données avec les utilisateurs des Centres d'accès aux données. Il existe différentes façons de partager des données confidentielles avec les utilisateurs qui ne sont pas assujettis à la *Loi sur la statistique*. La Loi comprend une disposition pour le partage des données avec les utilisateurs. Si l'on obtient la permission du répondant au moment de la collecte des données, des microdonnées confidentielles peuvent être partagées avec certains utilisateurs. Le partage des données serait possible avec le CRSH ou un « Institut » créé par cet organisme. Les répondants seraient informés de la nature des travaux de l'Institut (ou du CRSH), et on leur demanderait la permission de partager les données. Compte tenu des expériences passées, il est probable que la plupart des répondants accepteront. Dans un tel contexte, Statistique Canada fournirait alors les microdonnées brutes à l'Institut (ou au CRSH), lequel devrait évidemment convenir, en vertu d'un contrat, d'en assurer la confidentialité.

Cette approche a comme avantage qu'elle est explicite. Le processus de sélection et d'examen approfondi serait donc entièrement du ressort de l'Institut, et Statistique Canada jouerait un rôle minime, à savoir, établir les modalités nécessaires pour assurer la confidentialité des données. Cette approche a comme inconvénient qu'elle ne peut être appliquée qu'aux données recueillies à l'avenir, et à certains ensembles de données. Elle ne pourrait s'appliquer aux ensembles de données qui ont été créés jusqu'à maintenant. Cette approche devrait donc être sérieusement envisagée pour l'avenir, mais elle ne résoudra pas les problèmes actuels. Ainsi, il est possible que l'on constitue un « Institut » avec lequel on pourrait partager des données. Il faudrait vérifier dans quelle mesure les répondants sont disposés à partager des données financières avec un tel institut (ou le CRSH), aux fins de la recherche. Même si cette approche ne comportait que peu

d'avantages à court terme, il s'agirait d'une excellente solution à long terme pour le partage d'ensembles de données choisis.

Traitement des documents de recherche une fois le projet terminé. Il se peut que les chercheurs envisagent la production de deux documents une fois le projet terminé. Un rapport de recherche comprenant l'analyse quantitative et une interprétation des résultats serait remis au Comité de sélection et d'examen des Centres d'accès aux données de recherche. Ce document ferait partie de la série des documents de recherche[3] et ferait l'objet de l'examen habituel effectué par les pairs du milieu universitaire. Le Comité (ou son remplaçant désigné) gérerait ce processus. Le produit ferait aussi l'objet d'un examen « institutionnel » par Statistique Canada. Cet examen viserait uniquement à déterminer si les travaux sont prévus dans le mandat du Bureau relativement à la série des documents de recherche. Les travaux effectués pour la série des documents de recherche seraient généralement axés sur les politiques, mais ne pourraient contenir de recommandations directes en matière de politiques. Ainsi, l'examen institutionnel effectué par Statistique Canada servirait principalement à veiller à ce que le document ne comprenne pas de recommandations directes en matière de politiques[4]. Une fois les examens exécutés par les chercheurs et l'examen institutionnel terminés, les auteurs apporteraient les changements nécessaires.

Il va sans dire que les chercheurs seraient libres de publier le document de recherche, ou une version élargie de celui-ci, y compris des observations en matière de politiques et d'autres ajouts au besoin, dans le cadre de tout projet ou publication universitaire. En résumé, outre le document initial à caractère quantitatif soumis pour la série des documents de recherche, le chercheur serait libre de produire toute autre version du document et de la publier dans tout autre contexte. Advenant le cas où l'examen du document par le Comité de sélection et d'évaluation et Statistique Canada entraînait la décision de ne pas publier le document dans la série des documents de recherche, l'auteur pourrait le soumettre pour qu'il soit publié ailleurs[5].

6. Emplacement possible des centres

Les centres pourraient être affiliés à une université, un institut de recherche ou un réseau de recherche. Ils pourraient être situés dans les bureaux régionaux de Statistique Canada, et pour les raisons énoncées ci-après, il est proposé que l'on ait initialement recours aux bureaux régionaux à cette fin. Il va sans dire que l'emplacement physique doit être sécuritaire. Si on décidait d'utiliser des locaux à l'extérieur de Statistique Canada, il faudrait avoir recours à un processus concurrentiel pour en déterminer l'emplacement. Il serait préférable de commencer

---

[3] Il serait préférable de conserver dans un seul endroit la série des documents de recherche servant à la publication de l'analyse quantitative initiale. La gestion centrale de la série des documents de recherche permettrait d'assurer une certaine cohérence quant au traitement des documents. Cela pourrait entraîner des répercussions quant aux ressources pour Statistique Canada et le Comité. L'examen approfondi des documents de recherche, afin de s'assurer que les travaux sont correctement effectués, peut parfois se révéler long. Ces examens approfondis seraient effectués, pour une large part, dans le cadre du processus d'arbitrage universitaire. Il pourrait parfois arriver que des travaux soient nécessaires pour valider l'utilisation des données. Il est important que la personne responsable du maintien de la série comprenne bien le monde de la recherche et dispose des ressources nécessaires pour prévenir les arriérés de classement.

[4] Par recommandations ou observations directes en matière de politiques, on entend l'évaluation directe, la critique ou la promotion de programmes gouvernementaux existants ou proposés. Cela n'exclut pas les recherches sur des sujets pertinents aux politiques, qui sont en fait encouragées. Les restrictions s'appliquent aux produits de Statistique Canada (et à la série des documents de recherche conjoints), en vue d'en protéger la neutralité et l'objectivité.

[5] Normalement, les droits d'auteurs relatifs aux travaux qui ont été effectués à contrat, à titre de services spéciaux au ministre (aux termes de la Loi) appartiennent à Statistique Canada. Toutefois, il est proposé que, dans le cadre du contrat conclu avant le début des travaux, Statistique Canada convienne de remettre les droits d'auteurs aux chercheurs. Statistique Canada conserverait le droit d'examiner de façon approfondie toutes les publications découlant du projet, pour des raisons de confidentialité ou de fiabilité des données. Dans le cadre du même contrat, Statistique Canada conserverait le droit de reproduire le document au besoin.

avec un nombre très restreint de centres, étant donné que ce projet a des répercussions substantielles sur les ressources de Statistique Canada, du fait des services à fournir. Il semble probable que l'institut qui disposerait d'un tel centre pourrait servir de base pour l'établissement d'une capacité importante de recherche empirique, dans un certain nombre de disciplines. Des chercheurs très qualifiés qui se spécialisent dans la recherche empirique seraient attirés par un institut comportant un centre de données, ce qui permettrait à l'université ou à l'institut concerné de mettre en place un programme efficace[6].

7. Débuts modestes

Le programme proposé est ambitieux. Une approche pilote pourrait se révéler appropriée. Il est proposé que les ensembles de données mis initialement à la disposition des centres comprennent les nouvelles enquêtes longitudinales : Enquête sur la dynamique du travail et du revenu (et précédente Enquête sur l'activité), Enquête longitudinale nationale sur les enfants et les jeunes, Enquête nationale sur la santé de la population, Enquête sur le lieu de travail et les employés, et sources connexes de données pour les statistiques sociales nécessaires aux travaux de recherche. C'est là que la nécessité d'un accès direct aux microdonnées se fait le plus sentir. Il est aussi proposé que les bureaux régionaux servent d'emplacements initiaux pour le projet pilote. Cela permettra de procéder graduellement à l'établissement des centres. Néanmoins, l'accès aux microdonnées confidentielles serait grandement amélioré, et cela entraînerait une certaine équité quant à l'emplacement géographique des centres, étant donné que les bureaux régionaux sont disséminés partout au pays. S'ils étaient installés à Statistique Canada, les premiers centres pourraient être moins coûteux à faire fonctionner, étant donné qu'une partie de l'infrastructure existe déjà. Lorsque le projet pilote sera en œuvre depuis un certain temps (peut-être après deux ans), il est proposé de réexaminer le programme, et sous réserve de cet examen, de tenir un concours pour l'établissement de Centres d'accès aux données de recherche dans les universités ou d'autres organisations de recherche.

8. Financement

Les centres similaires exploités par le U.S. Census Bureau et la U.S. National Science Foundation disposent d'un budget annuel de 250 000 $ US par emplacement. Il existe au moins deux stratégies en matière de financement. La première s'apparente à celle relative à l'Initiative de démocratisation des données (IDD). En vertu de ce scénario, les centres disposeraient d'un financement global. Celui-ci serait assuré par le CRSH, l'université ou l'institut où le centre est situé (après l'étape pilote), et d'autres organisations dont les chercheurs utilisent les centres, par exemple, les ministères gouvernementaux, instituts de recherche, etc. Étant donné que l'on s'attend à ce que les centres soient principalement utilisés par des chercheurs universitaires, la majorité du financement global proviendrait du CRSH. Les chercheurs utilisant les centres n'auraient pas à assumer de coûts directs. Le processus de sélection dirigé par le Comité de sélection et d'examen permettrait de réglementer l'accès aux centres. Comme variante à cette approche, on pourrait compenser une part importante des coûts annuels par un financement global, et le reste par des droits versés par les chercheurs utilisant les centres. Les chercheurs universitaires dont les projets seraient approuvés par le Comité recevraient une subvention du CRSH (par l'entremise du Comité d'examen) pour compenser ces coûts. Les chercheurs d'autres organisations paieraient des droits similaires. Il est proposé de mettre en œuvre une de ces stratégies de financement.

---

[6] Il va sans dire que le NBER (National Bureau of Economic Research) aux États-Unis, l'un des premiers instituts disposant d'un centre de données pour le recensement, a mis sur pied un programme de recherche empirique exhaustif, peut-être en partie du fait que les chercheurs qui y sont affiliés ont accès à des données de qualité supérieure.

9. Conclusion

La présente proposition prévoit la prise en compte des Centres d'accès aux données de recherche comme l'une des solutions au problème de l'accès aux microdonnées pour les chercheurs. Initialement, il s'agirait de l'accès aux données des enquêtes auprès des ménages, particulièrement les nouvelles enquêtes longitudinales. Les enquêtes auprès des entreprises (établissements ou compagnies) posent des problèmes particuliers du point de vue de la confidentialité, qui pourraient empêcher qu'on les inclue dans ce projet, au moins initialement. Les centres pourraient être appuyés par une capacité d'accès à distance à l'intérieur de Statistique Canada. Cette dernière serait utile pour les chercheurs qui, du fait de leur situation géographique, ne peuvent utiliser un centre de données ou qui, pour d'autres raisons, ne peuvent ou ne veulent pas utiliser de tels centres. Les projets moins importants, notamment, recevraient un meilleur traitement grâce à l'accès à distance.

Annexe D

Développement d'une capacité d'accès à distance à Statistique Canada

Grâce à une capacité d'accès à distance, les chercheurs disposeraient des outils nécessaires pour leur permettre de préciser des procédures statistiques et de faire appliquer par le personnel de Statistique Canada ces procédures aux données confidentielles. Les résultats feraient l'objet d'un examen de la part de Statistique Canada, pour assurer la confidentialité, et seraient renvoyés aux chercheurs.

Parmi les outils mis à la disposition du chercheur figureraient le fichier de microdonnées à grande diffusion (dans la mesure du possible), et un fichier permuté, correspondant étroitement au fichier confidentiel, du point de vue de la structure et du niveau de détail, mais ne comportant qu'un nombre limité de données réelles. Les chercheurs préciseraient le programme ou le passage souhaité, sous forme de travaux SPSS/SAS, à partir du fichier permuté, et le soumettrait à Statistique Canada, par Internet. Les fichiers permutés devraient comprendre suffisamment de détails pour servir de base à la mise à l'essai et à la mise au point des programmes. Statistique Canada ne jouerait pas de rôle relativement à cette mise au point, et les erreurs de programmation seraient renvoyées aux chercheurs. La principale tâche des employés de Statistique Canada consisterait à exécuter les travaux à partir du fichier principal et à examiner le produit de façon approfondie, pour en assurer la confidentialité. Les programmes donnant lieu à des problèmes de confidentialité devraient être modifiés par le chercheur.

Les principaux avantages de l'accès à distance est la disponibilité des données pour les chercheurs, peu importe leur situation géographique, et la simplification de l'interface entre eux et Statistique Canada. Ses inconvénients sont les coûts (comparativement à l'utilisation directe de fichiers de données à grande diffusion) et les retards possibles découlant de l'examen approfondi des résultats. Un groupe de travail informel a été constitué à l'intérieur de Statistique Canada pour élaborer une stratégie de mise en œuvre qui réduirait ces inconvénients. Pour le moment, elle prend la forme suivante : les fichiers seraient produits par le secteur spécialisé responsable de l'enquête, avec l'aide du personnel de la méthodologie ayant une expérience de la production de tels fichiers. Ils seraient redistribués parmi les chercheurs, au moyen du site FTP de l'Initiative de démocratisation des données (IDD); les personnes-ressources de l'IDD dans les universités pourraient contribuer à faciliter l'accès aux fichiers, comme elles le font actuellement pour les fichiers de l'IDD. Statistique Canada établirait un groupe central, à partir de l'équipe actuelle de l'IDD, pour exécuter les programmes et vérifier les résultats aux fins de la confidentialité. Les résultats seraient renvoyés aux chercheurs au moyen du FTP.

Les principaux défis que comporte cette approche ont trait à la gamme des fichiers permutés qui peuvent être produits et fournis, ainsi qu'à l'efficacité de cette démarche, compte tenu des efforts nécessaires pour vérifier les résultats aux fins de la confidentialité, et du temps de traitement. Nombre de ces questions ne trouveront réponse qu'après une mise à l'essai appropriée de la stratégie. Cette dernière a été appliquée à l'Enquête nationale sur la santé de la population (ENSP), et on s'attend à ce que quatre fichiers soient disponibles pour la mise à l'essai dans un contexte d'accès à distance, l'an prochain. Cela devrait fournir aux chercheurs et à Statistique Canada suffisamment d'information pour décider comment utiliser l'accès à distance, de concert avec d'autres approches, en vue d'améliorer l'accès aux données.

Annexe E

Modèle de forums de recherche en statistiques sociales : le Forum canadien de recherche sur la situation d'emploi (FCRSE)

Le FCRSE est une société sans but lucratif dont l'objectif premier consiste à améliorer les interactions positives entre les chercheurs qui étudient les questions liées à l'emploi et les universités ainsi que les gouvernements et les décideurs eux-mêmes, de même qu'à accroître le volume et la qualité des recherches pertinentes aux politiques dans ce domaine. Le FCRSE est dirigé en alternance par un conseil bénévole, dont au moins le tiers provient du secteur universitaire et des services gouvernementaux. Le FCRSE reçoit un financement de base de Développement des ressources humaines Canada, mais il obtient aussi une quantité considérable des crédits pour la tenue de conférences de la part d'organismes de financement, de fondations et de ministères gouvernementaux concernés.

Au cours des huit dernières années, le FCRSE a organisé une série de conférences (près de vingt jusqu'à maintenant) sur une gamme variée de sujets de recherche. Cela a eu pour effet d'accroître de façon substantielle l'accès des décideurs à l'expertise technique des chercheurs universitaires, à un très faible coût. Le FCRSE a aussi suscité une augmentation considérable de la recherche à partir de données canadiennes. Cela n'est pas accidentel : les organisateurs de conférences du FCRSE ont encouragé la recherche au Canada, même s'ils reconnaissent les avantages des comparaisons avec les États-Unis et d'autres nations. Grâce à ses relations au gouvernement, le FCRSE a aussi facilité l'accès à des données confidentielles. Voici une liste des conférences organisées par le FCRSE.

- Conférence de fondation du FCRSE, Kingston (Ontario), 31 mai 1991.
- Recherche en matière de politiques sur la formation, le chômage et le soutien du revenu, ainsi que l'immigration, Aylmer (Québec), 6 et 7 mars 1992.
- **Le marché du travail au cours des deux dernières récessions : un point de vue comparatif (de concert avec Statistique Canada), Ottawa (Ontario), 5 mars 1993.**
- Programmes et politiques de soutien du revenu (de concert avec Santé et Bien-être social Canada), Ottawa (Ontario), 24 septembre 1993.
- Immigration et marché du travail, Hull (Québec), 11 mars 1994.
- Marchés du travail et soutien du revenu : accent sur la Colombie-Britannique (de concert avec les ministères des services sociaux et de la formation professionnelle de la Colombie-Britannique), Vancouver, (C.-B.), 25 mars 1994.
- Ateliers en parallèle avec le Congrès des sociétés savantes, Calgary (Alberta), 13 juin 1994.
- Adaptation des jeunes au marché du travail (de concert avec le ministère de l'éducation de la Colombie-Britannique et l'UBC), Vancouver (C.-B.), 25 juin 1994.
- Conférence internationale sur l'évaluation de l'assurance-emploi, Ottawa (Ontario), 14 et 15 octobre 1994.
- Politiques à l'intention des travailleurs déplacés (de concert avec l'Université du Québec à Montréal et Ekos Research Associates), Montréal (Québec), 2 et 3 décembre 1994.
- Recyclage des travailleurs occupés : accent sur le Nouveau-Brunswick (de concert avec DRHC, région du Nouveau-Brunswick, et l'UNB), Fredericton (N.-B.), 31 mars et 1er avril 1995.
- Politiques environnementales et marchés du travail (de concert avec le secteur des Innovations de DRHC), Ottawa (Ontario), 15 et 16 septembre 1995.
- Écart entre les taux de chômage au Canada et aux États-Unis (de concert avec le Centre d'étude des niveaux de vie), Ottawa (Ontario), 9 et 10 février 1996.

- Changements touchant le temps de travail au Canada et aux États-Unis (de concert avec le Upjohn Institute, Statistique Canada, DRHC), Ottawa (Ontario), 13 au 15 juin 1996.
- Expériences sociales, évaluation, emploi et politique sociale, 1996.
- Immigration, marché du travail et économie, Richmond (C.-B.), 17 et 18 octobre 1997.
- Taux de participation et ratios d'emploi, Ottawa (Ontario), 17 décembre 1997.
- Transitions sur le marché du travail et dynamique du revenu (coparrainée par Statistique Canada et DRHC), parallèlement aux réunions de l'ACE à Ottawa (Ontario), 28 et 29 mai 1998.
- Conférence sur l'emploi autonome (de concert avec le Réseau international du travail au Canada et l'OCDE), Burlington (Ontario), 24 au 26 septembre 1998.

On peut obtenir plus de renseignements sur le FCRSE dans le site Web de l'organisme : http://cerf.mcmaster.ca

Annexe F

Composition du Groupe de travail mixte

Paul Bernard, président
Département de sociologie
Université de Montréal

Peter Kuhn
Department of Economics
McMaster University

Douglas A. Norris, directeur
Division de la statistique sociale, du logement et des familles
Statistique Canada

Garnett Picot, directeur
Division de l'analyse des entreprises et du marché du travail
Statistique Canada

Martin Wilk, ancien statisticien en chef
Statistique Canada

Betty Havens
Community Health Sciences
University of Manitoba

Céline Le Bourdais
INRS – Urbanisation

Michael Ornstein
Institute for Social Research
York University

J. Douglas Willms
Atlantic Centre for Policy Research in Education
University of New Brunswick

Aidés par :

Hélène Régnier, analyste des politiques
CRSH